The Child Care and Early Education Heart Monitor: The Intersection of Structural Quality and Process Quality Using the Contact Hour Metric Proposal

Richard Fiene PhD

Penn State Edna Bennett Pierce Prevention Research Center

June 2025

The Child Care and Early Education (CCEE) field needs a means to monitor the key elements of structural and process quality in a unified means. The theory of regulatory compliance has been suggested as a unifying framework for structural and process quality; but at a more practical level what could be used to essentially unify the monitoring and measurement of both structural and process quality. Generally, structural and process quality are measured separately from each other by using very separate and distinct tools utilized by licensing inspectors and quality observers. This research abstract will build off several papers (see Appendix) that deal with the creation of a new Contact Hour (CH) metric replacing measuring compliance with adult-child ratios, unifying structural quality with process quality, and dealing with the results from the uncertainty-certainty matrix as it relates to reliability and validity in CCEE.

Let's begin by revisiting the Contact Hour (CH) metric. It will be summarized in this introductory section with the original paper attached in the Appendix. The same will be done with the other two concepts dealt with on unifying structural and process quality, and the uncertainty-certainty matrix. The Contact Hour metric has been proposed as a more effective and efficient metric for measuring compliance with adult-child ratios and group sizes in CCEE programs. It is simple to apply by just asking 6 questions about when children arrive and leave a CCEE program and how many staff are present in a particular classroom. Once that is done a trapezoidal model is built in which compliance with staff child and group size rules can be determined. Please refer to the Contact Hour papers in the Appendix which provides the mathematical details on how this is done. Also, the Contact Hour papers provide some additional insights in how the Contact Hour metric can be used for other preventive purposes - please refer to the second Contact Hour 2021 paper in the Appendix..

In the results, the contact hours are dealt with as absolute values but let's enhance this result by moving it from an absolute value to one that is more relative by introducing process quality measures such as from the Environmental Rating Scales (ERS) or the Classroom Assessment Scoring System (CLASS). To do this, it would take 1000's of observations to fill the contact hour trapezoidal model which is not realistic. But let's let Artificial Intelligence (AI) do the observing and training of AI in what constitutes the various quality levels on the respective ERS and CLASS tools. By using AI and having video cameras in each of the classrooms to be assessed, this becomes doable. The ERS AI Observer or the CLASS AI Observer would be able to collect the data by observing and assessing what it sees via the video cameras installed in the classrooms. Summary measurements would be made on an hourly basis and recorded as part of the Contact Hour trapezoidal model. At the end of the day, there would be a relative value utilized in this model rather than the absolute value that has been used in the past to determine structural quality compliance with adult-child ratio and group size. For example, if a CCEE program classroom exceeded the area of the trapezoidal model it was out of compliance and if it were within the area of the trapezoidal model it was in compliance. By adding the ERS and/or CLASS AI data, it changes this metric totally by adding process quality measures which can be measured on a 1-7 ordinal scale (see the Development of a Regulatory Compliance Scale paper in the Appendix for additional details about this approach).

This approach will get at the heart of CCEE monitoring, process quality, measuring the interactions amongst staff and children in an ongoing fashion. It moves the needle from being structural to process quality providing an intersection of both components of quality. The AI approach will also help to address the issues related to bias in regulatory compliance observing and decision making by inspectors/observers. By training the AI ERS and CLASS Observers there should be greater certainty established in making the right decisions related to specific quality elements. Just as in establishing inter-rater reliability with human observers, the same can be done with the ERS and CLASS AI Observers. Please refer to the Uncertainty-Certainty Matrix paper, and the Structural and Process Quality paper in the appendix.

Appendices Follow:

Contact Hour 2020 Paper

Contact Hour 2021 Paper

Uncertainty-Certainty Matrix 2024 Paper

Structural and Process Quality 2025 Paper

Development of a Regulatory Compliance Scale 2025 Paper

Contact Hour Pilot Study Design Proposal
Richard Fiene, Ph.D.
Research Institute for Key Indicators
April 2020

The purpose of this proposal is to develop the key parameters for testing out the Contact Hour (CH) methodology in a series of facilities to determine its efficacy. The pilot will determine if this CH methodology has any merit in being able to be used as a rough estimate to identifying facilities that may be at greater risk to spreading an infectious disease, such as the COVID19 virus. Since monitoring of facilities will not be occurring during the COVID19 pandemic are there ways to measure the research question in the previous sentence. Yes there is and it is based upon the Contact Hour (CH) methodology and involves asking the following seven questions[1]:

1. When does your first teaching staff arrive or when does your facility open?
2. When does your last teaching staff leave or when does your facility close?
3. Number of teaching/caregiving staff?
3. Number of children on your maximum enrollment day?
5. When does your last child arrive?
6. When does your first child leave?
4. Has any child or adult within your facility contracted the COVID19 virus?

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

(1) CH=((NC(TO+TH))/2)/TA; (2) CH=(NCxTO)/TA; (3) CH=((NCxTO)/2)/TA; (4) CH=(NC$^{2)/TA}$

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

By knowing the number of contact hours (CH) it will be possible to rank order the exposure time of adults with children. This metric could then be used to determine if greater contact hours is correlated with the increased risk of the COVID19 virus. The COVID19 virus question is the dependent variable and is not used in the above formulae.

The following chart can be used by entering the following metrics (example in the table is based upon 5 enrolled children (NC)): the facility is open for 10 hours (TO) and then various scenarios are played out for how long the facility is at full enrollment (TH). Based upon these metrics an outcome rubric can be used where less CH is a positive (+), while high CH is a negative (-). For simplicity, the following chart is based upon one teaching staff (TA) being present (1:5 Adult-Child Ratio). The chart on page 2 provides a more detailed depiction of various CH for a multitude of Adult-Child Ratios and the figure on page 3 shows a hypothesized relationship between CH and COVID19 infection rates.

Contact Hour Score Generated from Above 4 Formulae and Potential Outcomes (COVID19 Infections)

Contact Hours - CH Score	Formulae for CH Score	Potential Outcomes
10	(2 (NC) x 10 (TO)) / 2	+ (None or few COVID19 Infections)
38	(5 (NC) (5 (TH) + 10 (TO)) / 2	+ / - (Lower # of COVID19 Infections)
80	8 (NC) x 10 (TO)	-/ + (Higher # of COVID19 Infections)
100	10 (NC) x 10 (TO)	-(Highest # of COVID19 Infections)

Contact Hour (CH) Conversion Table (Fiene, 2020©)

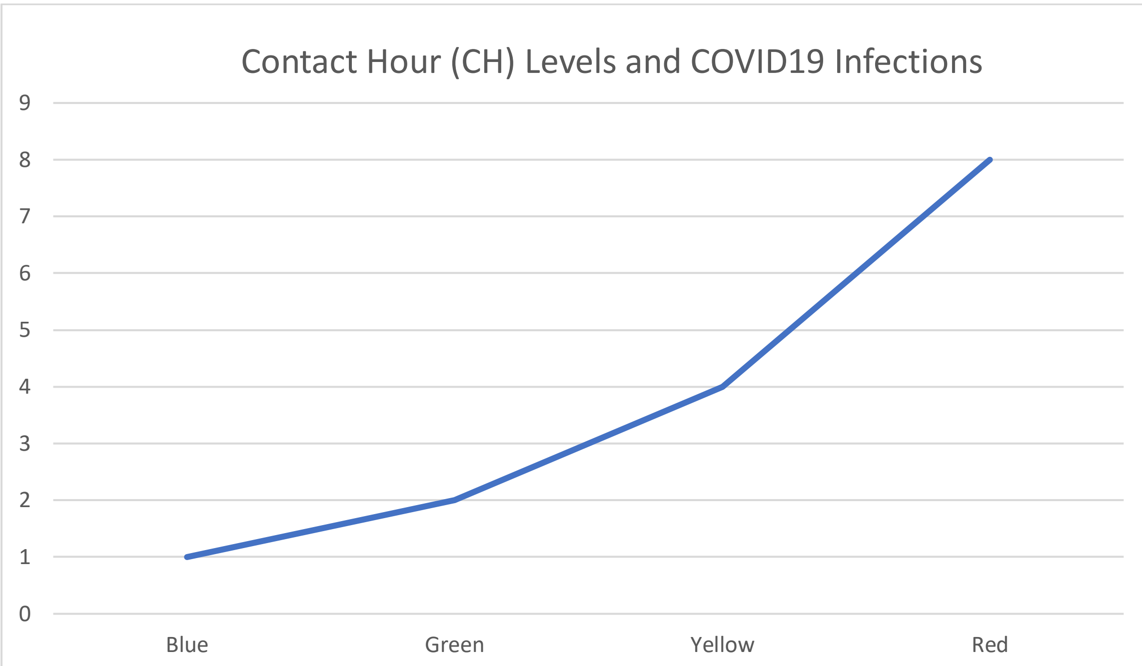
The previous chart on page 1 provided a theoretical view of how Contact Hours could be calculated, the following chart provides the addition of the number of staff (TA) in the equation and enhances the Contact Hours metric by calculating a Relatively Weighted Contact Hours (RWCH).

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

NC	CH GS	1:1	2:1	3:1	4:1	5:1	6:1	7:1	8:1	9:1	10:1	11:1	12:1	13:1	14:1	15:1
1	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
2	20	10	20	20	20	20	20	20	20	20	20	20	20	20	20	20
3	30	10	15	30	30	30	30	30	30	30	30	30	30	30	30	30
4	40	10	20	20	40	40	40	40	40	40	40	40	40	40	40	40
5	50	10	17	25	25	50	50	50	50	50	50	50	50	50	50	50
6	60	10	20	30	30	30	60	60	60	60	60	60	60	60	60	60
7	70	10	18	23	35	35	35	70	70	70	70	70	70	70	70	70
8	80	10	20	27	40	40	40	40	80	80	80	80	80	80	80	80
9	90	10	18	30	30	45	45	45	45	90	90	90	90	90	90	90
10	100	10	20	25	33	50	50	50	50	50	100	100	100	100	100	100
11	110	10	22	28	37	37	55	55	55	55	55	110	110	110	110	110
12	120	10	20	30	40	40	60	60	60	60	60	60	120	120	120	120
13	130	10	22	26	33	43	43	65	65	65	65	65	65	130	130	130
14	140	10	20	28	35	47	47	70	70	70	70	70	70	70	140	140
15	150	10	21	30	38	50	50	50	75	75	75	75	75	75	75	150
16	160	10	20	27	40	40	53	53	80	80	80	80	80	80	80	80
17	170	10	21	28	34	43	57	57	57	85	85	85	85	85	85	85
18	180	10	20	30	36	45	60	60	60	90	90	90	90	90	90	90
19	190	10	21	27	38	48	48	63	63	63	95	95	95	95	95	95
20	200	10	20	29	40	50	50	67	67	67	100	100	100	100	100	100
21	210	10	21	30	35	42	53	70	70	70	70	105	105	105	105	105
22	220	10	20	28	37	44	55	55	73	73	73	110	110	110	110	110
23	230	10	21	29	38	46	58	58	77	77	77	77	115	115	115	115
24	240	10	20	30	40	48	60	60	80	80	80	80	120	120	120	120
25	250	10	21	28	36	50	50	63	63	83	83	83	83	125	125	125
26	260	10	20	29	37	43	52	65	65	87	87	87	87	130	130	130
27	270	10	21	30	39	45	54	68	68	90	90	90	90	90	135	135
28	280	10	20	28	40	47	56	70	70	70	93	93	93	93	140	140
29	290	10	21	29	36	48	58	58	73	73	97	97	97	97	97	145
30	300	10	20	30	38	50	60	60	75	75	75	100	100	100	100	150

This table is based upon the assumptions that the child care is 10 hours in length (TO) and that the full enrollment is present for the full 10 hours (TH). This is unlikely to ever occur but it gives us a reference point to measure adult child contact hours in the most efficient manner. Based upon the relationship between TO and TH, select from one of the formulae from the previous page (1-4) to determine how well the actual Relatively Weighted Contact Hours (RWCH) match with this table. If the RWCH exceed the respective RWCH in this table, then the facility would be over ratio on ACR and exceed group size standards.

Based upon the above tables classifications, the following figure provides a hypothesized relationship between the various contact hour (CH) levels of blue, green, yellow, and red and the ranges these color schemes represent as per COVID19 infections.



The above figure's hypothesized results projects that as the level of Contact Hours (CH) increases, a corresponding increase in COVID19 infections in adults and children would also occur in the child care facility starting off slowly at the lowest level of CH (Blue), increasing slightly (Green), but then a steeper curve (Yellow), and steepest at the Red level where CH would be at the highest representing the greatest number of children and adults present over time.

The proposed pilot study will test this hypothesis to determine if this is the case or not[2].

Notes:

1 The seven (7) questions should be asked of each grouping that is defined by a classroom or a well defined group within each classroom tied to a specific adult-child ratio.

2 The results from this pilot study could lead to interesting planning for the future in which a particular threshold could be identified where the infection rates are too high or where infection rates begin.

Additional information regarding this methodology can be obtained from contacting: Dr Richard Fiene, Research Psychologist, Research Institute for Key Indicators, & Penn State University. RFiene@RIKInstitute.com or RJF8@psu.edu. http://RIKInstitute.com

Contact Hours as a New Metric Replacing Group Size and Staff-Child Ratios as well as a New Metric for COVID19 Thresholds

Richard Fiene, Ph.D.
Research Institute for Key Indicators
Sonya Stevens, Ed.D.
Washington Department of Children, Youth and Families

May 2021

The purpose of this paper is to propose Contact Hours as a new metric replacing staff child ratios and group size as well as using it as a new threshold measure for COVID19 thresholds. This paper will attempt to validate the key parameters for testing out the Contact Hour (CH) methodology in a series of facilities to determine its efficacy. The pilot validation study will determine if this CH methodology has any merit in being able to measure regulatory compliance with adult-child ratios. Since monitoring of facilities will not be occurring during the COVID19 pandemic are there ways to measure the research question in the previous sentence. Yes there is and it is based upon the Contact Hour (CH) methodology and involves asking the following six questions (The six questions should be asked of each grouping that is defined by a classroom or a well-defined group within each classroom tied to a specific adult-child ratio.):

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

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

$$(1)\ CH = ((NC\ (TO + TH)) / 2) / TA;$$

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

$$(3)\ CH = ((NC \times TO) / 2) / TA;$$

$$(4)\ CH = (NC^2) / TA$$

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

By knowing the number of contact hours (CH) it will be possible to rank order the exposure time of adults with children. Theoretically, this metric could then be used to determine that the greater contact hours is correlated with the increased non-regulatory compliance with adult-child ratios as determined in the below table on page 2.

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

Taking into Account Exposure Time and Density

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

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

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

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

(RS Model = 1.0)

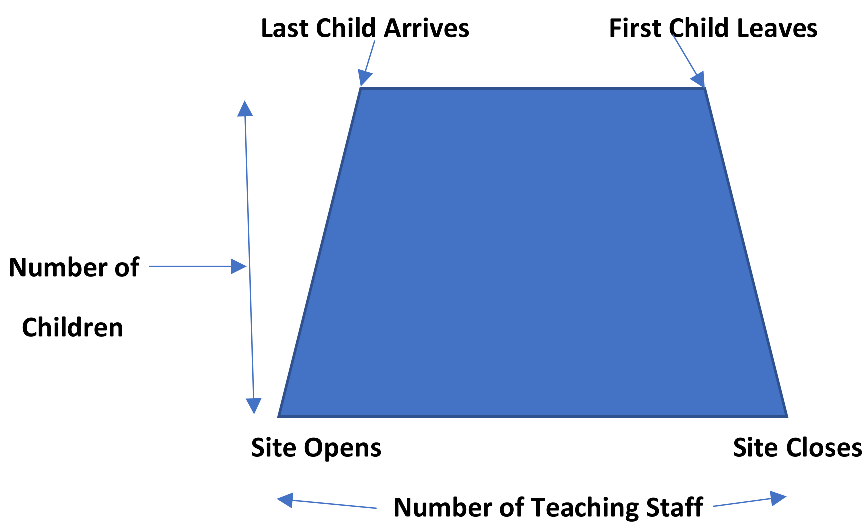
(TT Model = 0.5)

Sample/Data Collection Methods

Child care attendance data was explored and collected in partnership with the Washington State Department of Children, Youth, and Families (DCYF). A convenient sample of center and school age providers was initially identified through the use the state subsidy electronic payment system. All providers who accept Working Connections Child Care subsidies are required to use and track child attendance using an electronic attendance system. Providers may use and electronic sign in and out system provided by the state or opt to use another system. For this validation process, the sample was identified from the attendance tracking system provided and operated by DCYF and was inclusive of providers who use the system to track attendance of both subsidy and private pay children. The search resulted in approximately 100 providers within the State of Washington who have opted to use the electronic check-in system for all children regardless of payment type.

The sample was prioritized by identifying a single week since the Covid-19 outbreak began and from there the highest attendance day for that week was chosen for each provider. From this narrowed data set, it was determined the exact time the last child for the chosen day checked in, when the first child left, how many children were in attendance that day and the regular operating hours of the center or school age program. Because the attendance tracking system does not also track staffing attendance, it was necessary to contact each provider by phone in order to gather data inclusive of when the first staff arrived and when the last staff left and the total staff working that day. All responses were voluntary. Additionally, providers confirmed operating hours (many had been temporarily adjusted due to lowered demand during the gubernatorial stay at home order). Finally, providers reported if a child or staff member had tested positive for Covid-19. Of the 100 phone calls, the final sample was inclusive of 88 licensed providers statewide. Twelve providers either did not answer the call or opted to not answer the questions.

<u>Figure 1: Contact Hour Diagram Paradigm and Schematic</u>



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

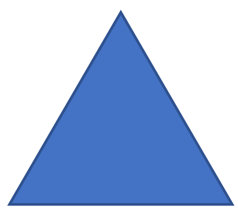
Potential Density Distributions
Taking into Account Number of Children, Staff, and Exposure Time

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

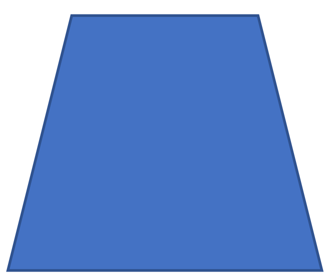
- RWCH = ACR
- CH = GS = NC
- NC and CH are highly correlated
- ACR and GS are static, not dynamic
- CH makes them dynamic by making them 2-D by adding in Time (T)
- ΣACR = GS
- GS = total number of children NC
- ACR = children / adult

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

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



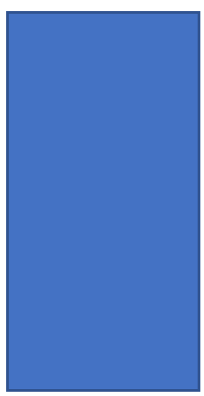
This density distribution should result in the lowest CH but probably not very likely to occur. Essentially what would happen is that full enrollment would be a single point which means that the last child arrives when the first child is leaving. Very unlikely but possible. (TT Model Reference(0.5))



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

This scenario is unlikely but is used as the reference point for CH because it provides the most efficient model. This is where all the children arrive and leave at the same time. Very unlikely, but I guess it could happen. The important element here is its efficiency in that all contact hours are covered, so although a lesser amount of CH is not as efficient it does demonstrate compliance with ACR and GS which is one of the purposes of CH. As the bottom two distributions will demonstrate, CHs above this level would either depict a program that is open for an extended time or where there are too many children present and the facility is out of compliance with GS and/or ACR. (RS Model Reference(1.0))

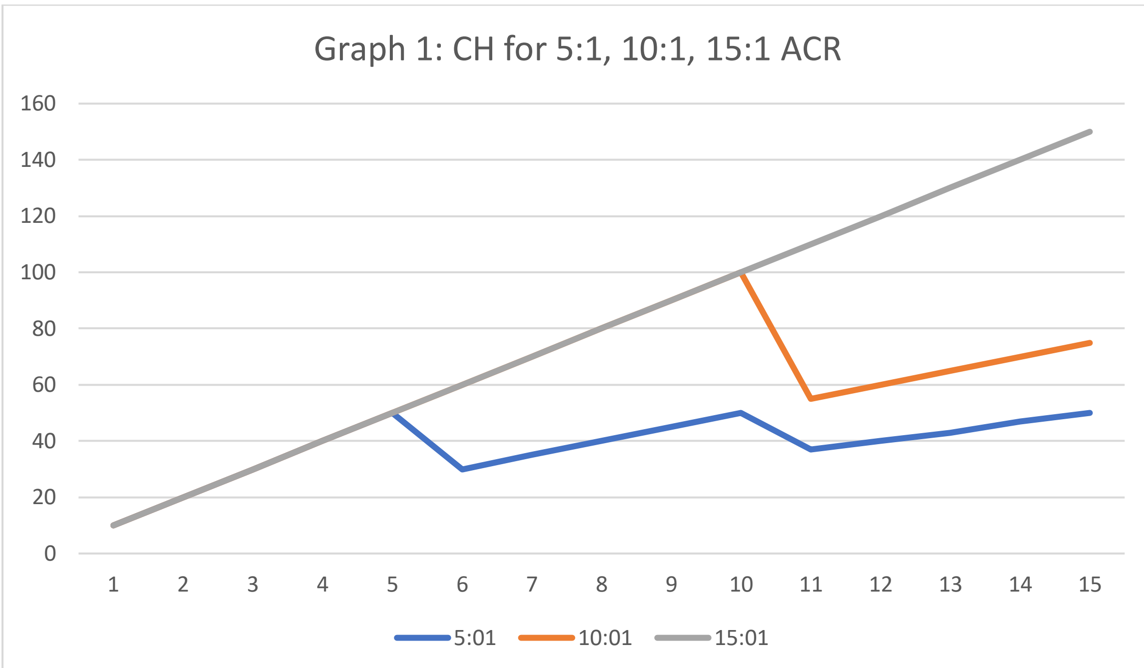
This distribution would indicate that the facility is open for an extended time and exceeds the number of total CH as depicted in the reference square standard. Although not out of compliance with GS or ACR, this could become a determining factor when looking at the potential overall exposure of adults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



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

There is some overlap in the RWCH (Table 1 on page 2) in moving across the various levels, that occurs because of the change in group size (GS) where an overall group size (GS) could influence the overall CH by increasing NC.

The below graph (Graph 1) depicts the contact hours (CH) for three different adult to child ratios (ACR) 5:1, 10:1 and 15:1 to demonstrate the relationship between CH & ACR as the number of children (NC) increases. CH is along the vertical axis, with NC along the horizontal axis.

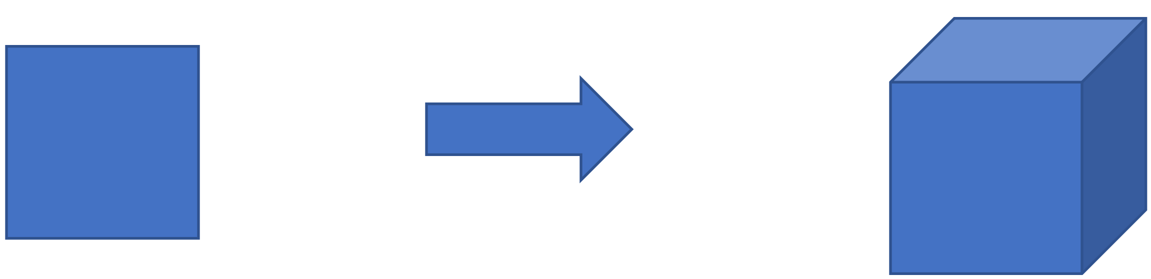


This graphic (Graph 1) depicts how with the addition of staff, the CH drop off accordingly.

****

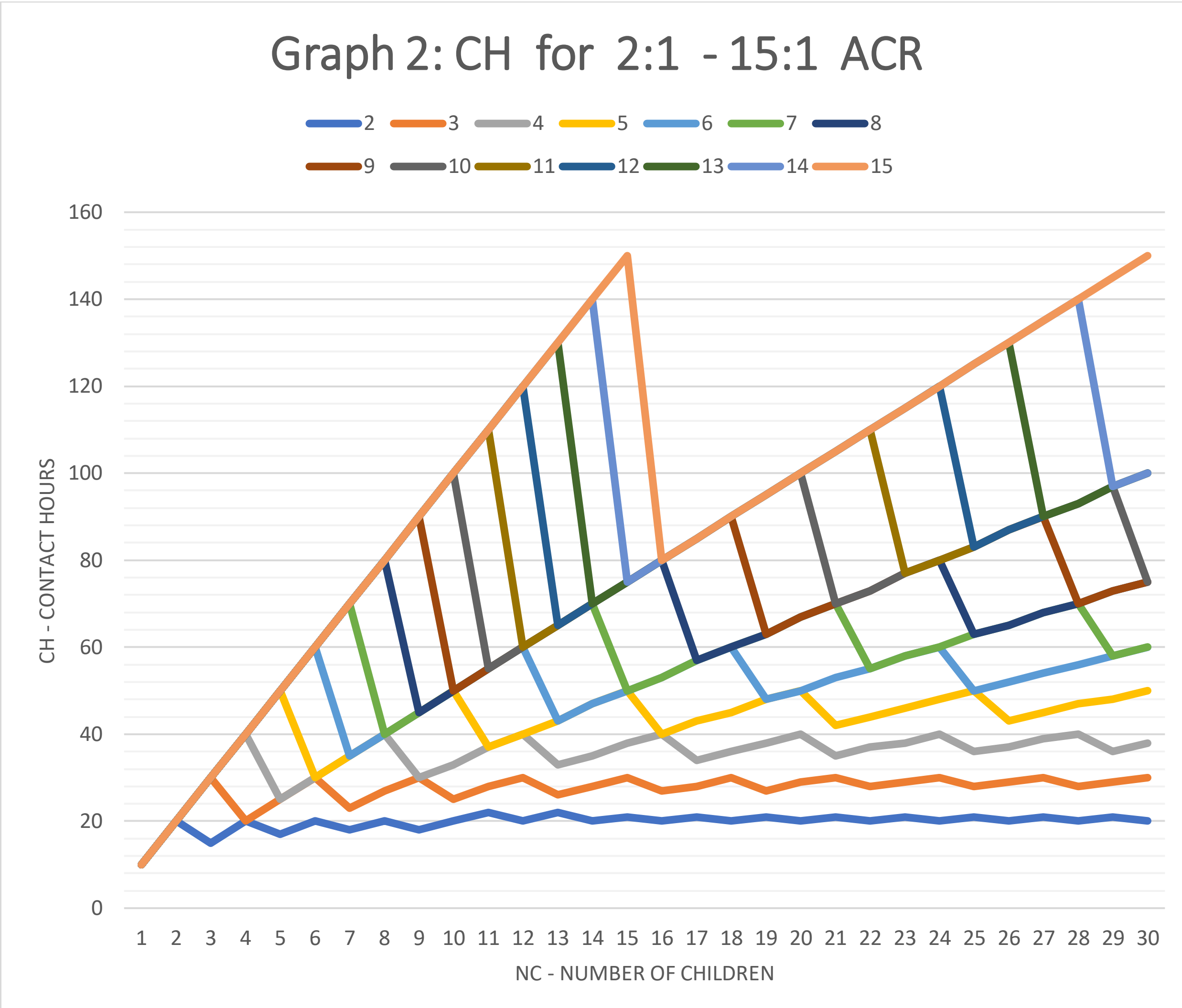
A possible extension or the next level to the CH methodology is to move from 2-D to 3-D and make the CH block format rather than area format. It could be used to describe the trilemma of accessibility, affordability and quality more fully. It could be a means for determining the unit cost at a much finer level and could then be used to make more informed decisions about the real cost of services.

Or another way of moving to 3-D is to include the square footage of the classroom or facility which would then provide a space metric along with time exposure and density metrics.



The move from 2-D (GS, ACR) to 3-D (GS, ACR, Quality or SQFT) and its potential impacts on the density distributions. Utilizing SQFT as a distancing/space dimension does help to mitigate the increased CH.

The following graph (Graph 2) depicts the Contact Hours (CH) for all the various Adult-Child ratios (ACR) in the Table on page 2 of this paper and how CH change with the number of children (NC).



From the above graph (Graph 2) it clearly shows how CHs vary with the number of children present. Please note the various slopes of the respective lines for each of the ACRs. As can be seen, once the lines begin to fluctuate, the CHs are entering into a zone of higher rate of exposure based on the ACRs. This demonstrates that the lower the ratio the more stable the CH line.

This is a listing of the algorithms for determining which formula (1-4 from page 1) & which model (RS or TT) to use in order to calculate the Contact Hours (CH). NC = Number of Children; TO = Total number of hours facility is open; TH = Total number of hours at full enrollment; TA = Total number of adult staff:

If TO = TH = NC, then (NC x TO)/TA = CH   (RS Model)

If TH < TO, then ((NC (TO + TH))/2)/TA = CH; or If TH = 0, then ((NC x TO)/2)/TA = CH (TT Model)

If TO = TH < NC, then (NC x TH)/TA = CH   (RS Model)

If TO = TH > NC, then (NC x TO)/TA = CH (RS Model)

Based upon the Washington State data, the Contact Hour methodology was validated in being able to act as a screener with those programs that would have exceeded the required staff child ratios. As can be seen through the data the more contact hours a staff person has with more children increases the probability of infection rates; when educators spend less time with lower amounts of children there is a lower chance of infection and vice versa. These data demonstrate how this methodology was used to assist in predicting appropriate child to adult ratios during an outbreak or pandemic by identifying safety thresholds of adult child ratios in licensed early learning facilities. The following spreadsheet plays out several scenarios with the actual data from Washington State early learning sites. For individuals interested in using the below spreadsheet in their respective jurisdiction, please contact the authors for the actual templates[1].

This provides evidence to support the use of this methodology in determining staff child ratio virtually as well as identifying when those ratios allow for in-person inspections or indicate when it is more appropriate to conduct virtual inspections. The authors do want to caution licensing administrators in that the results from this methodology is not to substitute for on-site observations when they are possible. It is intended as a screening tool to determine in a very overarching way how to target limited observational visits. The methodology is based upon statistical probabilities which have demonstrated in this pilot study to be highly reliable and valid but they are not full proof. So with any programs where there is any doubt, the agency should follow up with a direct observational inspection. Finally, agencies may want to consider using medical and geographical outbreak data in conjunction with this methodology to refine the results given the unique nature of the various infectious diseases.

In using the actual data from Washington State in the following spreadsheet, please note that the potential spread of the virus is mitigated the most greatly in the results in Green while Yellow and Red provide less mitigation and begin to place the adults and children at greater risk. Examples are provided for both the RS (1.0) and TT (0.5) Models

As a footnote to this study, a follow-up is to introduce distance/spacing via square footage (SQFT) to the Contact Hour formula. The results indicate a significant mitigation effect on increased Contact Hours when the available square footage is increased. This addition will be used in future studies to ascertain its relative impact on the Contact Hour formulas as indicated in the following revision.

CH2 = (((NC (TO + TH)) / 2) / TA) / (SQFT);

CH2 = ((NC x TO) / TA) / (SQFT);

CH2 = (((NC x TO) / 2) / TA) / (SQFT);

CH2 = ((NC^2) / TA) / (SQFT)

---

[1] Richard Fiene, Ph.D., Research Psychologist, Research Institute for Key Indicators and Affiliate Professor, Prevention Research Center, Penn State University. rjf8@psu.edu; http://prevention.psu.edu/people/fiene-richard

Sonya Stevens, Ed.D., Research Manager, Washington Department of Children, Youth, and Families, Olympia, Washington. Sonya.Stevens@dcyf.wa..gov

Contact Hour Models

NC	TA	TO	TH	TO+TH	(TOTH)N	CH	*RWCH*	RS Model - ACRCH 5:01	10:01	15:01
10	2	8	8	16	160	80	*40*	40	80	80
20	1	12	8	20	400	200	*200*	40	80	80
30	1	12	7	19	570	285	*285*	40	80	120
5	1	8	8	16	80	40	*40*	40	40	40
15	2	8	8	16	240	120	*60*	40	60	120
9	2	12	9	21	189	94.5	*47.25*	40	80	107

NC	TA	TO	NCxTO	CH	*RWCH*	TT Model - ACRCH 5:01	10:01	15:01
10	2	8	80	40	*20*	20	40	40
20	1	8	160	80	*80*	20	40	40
30	3	8	240	120	*40*	20	40	60
5	1	8	40	20	*20*	20	20	20
15	2	8	120	60	*30*	20	30	60
25	1	8	200	100	*100*	20	33	52

TT Model = CH=((NC(TO+TH))/2)/TA=RWCH; CH=((NCxTO)/2)/TA=RWCH; if TH<TO or if TH=0

RS Model = CH=(NCxTO)/TA=RWCH; CH=(NC2)/TA=RWCH; if TO=TH=NC or if TO=TH<NC or if TO=TH>NC

Legend: ***NC = Number of Children in attendance***
TA = Number of Teaching Staff
TO = Number of hours site is open
TH = Number of hours site at full enrollment
CH = Contact Hours with Children
RWCH = Relatively Weighted Contact Hours with Staff

Decisions:
Green = meets or under with all ACRs
Yellow = meets, under and over with ACRs
Red = over with all ACRs

Questions:
1) Number of children in attendance on your maximum enrollment day (NC)?
2) Number of teaching/caregiving staff on that same maximum enrollment day (TA)?
3) When does your first teaching staff arrive or when does your facility open (TO1)? **TO = TO2-TO1**
4) When does you last teaching staff leave or when does your facility close (TO2)?
5) When does you last child arrive (TH1)? **TH = TH2-TH1**
6) When does your first child leave (TH2)?

Table of Conversions - RS Model - ACRCH (Relatively Weighted Contact Hours)

NC	CH	1:01	2:01	3:01	4:01	5:01	6:01	7:01	8:01	9:01	10:01	11:01	12:01	13:01	14:01	15:01
1	8	**8**	8	8	8	8	8	8	8	8	8	8	8	8	8	8
2	16	**8**	**16**	16	16	16	16	16	16	16	16	16	16	16	16	16
3	24	**8**	12	**24**	24	24	24	24	24	24	24	24	24	24	24	24
4	32	**8**	**16**	16	**32**	32	32	32	32	32	32	32	32	32	32	32
5	40	**8**	13	20	20	**40**	40	40	40	40	40	40	40	40	40	40
6	48	**8**	**16**	**24**	24	24	**48**	48	48	48	48	48	48	48	48	48
7	56	**8**	14	19	28	28	28	**56**	56	56	56	56	56	56	56	56
8	64	**8**	**16**	21	**32**	32	32	32	**64**	64	64	64	64	64	64	64
9	72	**8**	14	**24**	24	36	36	36	36	**72**	72	72	72	72	72	72
10	80	**8**	**16**	20	27	**40**	40	40	40	40	**80**	80	80	80	80	80
11	88	**8**	15	22	29	29	44	44	44	44	44	**88**	88	88	88	88
12	96	**8**	**16**	**24**	**32**	32	**48**	48	48	48	48	48	**96**	96	96	96
13	104	**8**	15	21	26	35	35	52	52	52	52	52	52	**104**	104	104
14	112	**8**	**16**	22	28	37	37	**56**	56	56	56	56	56	56	**112**	112
15	120	**8**	15	**24**	30	**40**	40	40	60	60	60	60	60	60	60	**120**
16	128	**8**	**16**	21	**32**	32	43	43	**64**	64	64	64	64	64	64	64
17	136	**8**	15	23	27	34	45	45	45	68	68	68	68	68	68	68
18	144	**8**	**16**	**24**	29	36	**48**	48	48	**72**	72	72	72	72	72	72
19	152	**8**	15	22	30	38	38	51	51	51	76	76	76	76	76	76
20	160	**8**	**16**	23	**32**	**40**	40	53	53	53	**80**	80	80	80	80	80
21	168	**8**	15	**24**	28	34	42	**56**	56	56	56	84	84	84	84	84
22	176	**8**	**16**	22	29	35	44	44	59	59	59	**88**	88	88	88	88
23	184	**8**	15	23	31	37	46	46	61	61	61	61	92	92	92	92
24	192	**8**	**16**	**24**	**32**	38	**48**	48	**64**	64	64	64	**96**	96	96	96
25	200	**8**	15	22	29	**40**	40	50	50	67	67	67	67	100	100	100
26	208	**8**	**16**	23	30	35	42	52	52	69	69	69	69	**104**	104	104
27	216	**8**	15	**24**	31	36	43	54	54	**72**	72	72	72	72	108	108
28	224	**8**	**16**	22	**32**	37	45	**56**	56	56	75	75	75	75	**112**	112
29	232	**8**	15	23	29	39	46	46	58	58	77	77	77	77	77	116
30	240	**8**	**16**	**24**	30	**40**	**48**	48	60	60	80	80	80	80	80	**120**

RS Model = 1.0
TT Model = 0.5

The above examples are drawn from a pilot study done with Washington DCYF ECE facilities.

The Uncertainty-Certainty Matrix Logic Model and Algorithms

Richard Fiene PhD

Penn State Edna Bennett Pierce Prevention Research Center

May 2024

This research abstract will take the Confusion Matrix which is a well-known metric in the decision-making research literature and refocus it on regulatory science within the context of the definition of regulatory compliance and licensing measurement. It will also deal with the policy implications of this particular metric. In this abstract, it is proposed that the Uncertainty-Certainty Matrix (UCM) is a fundamental building block to licensing decision making. The 2 x 2 matrix is the center piece for determining key indicator rules, but it is also a core conceptual framework in licensing measurement and ultimately in program monitoring and reviews.

The reason for selecting this matrix is the nature of licensing data, it is binary or nominal in measurement. Either a rule/regulation is in compliance or out of compliance. Presently most jurisdictions deal with regulatory compliance measurement in this nominal level or binary level. There is to be no gray area, this is a clear distinction in making a licensing decision about regulatory compliance. The UCM also takes the concept of Inter-Rater Reliability (IRR) a step further in introducing an uncertainty dimension that is very important in licensing decision making which is not as critical when calculating IRR. It is moving from an individual metric to a group metric (See Figures 1 & 2) involving regulatory compliance with rules.

The key pieces to the UCM are the following: the decision (D) regarding regulatory compliance and actual state (S) of regulatory compliance. Plus (+) = In-compliance or Minus (-) = Out of compliance. So, let's build the matrix:

<u>Table 1: Uncertainty-Certainty Matrix (UCM) Logic Model</u>

UCM Matrix Logic		**Decision (D) Regarding**	**Regulatory Compliance**
		(+) In Compliance	(-) Not In Compliance
Actual State (S) of	(+) In Compliance	Agreement	Disagreement
Compliance	(-) Not In Compliance	Disagreement	Agreement

The above UCM matrix demonstrates when agreement and disagreement occur which establishes a level of certainty (Agreement Cells) or uncertainty (Disagreement Cells). In a perfect world, there would only be agreements and no disagreements between the decisions made about regulatory compliance and the actual state of regulatory compliance. But from experience, this is not the case based upon reliability testing done in the licensing research field in which a decision is made regarding regulatory compliance with a specific rule or regulation and then that is verified by a second observer who generally is considered the measurement standard.

Disagreements raise concerns in general, but the disagreements are of two types: false positives and false negatives. A false positive is when a decision is made that a rule/regulation is out of compliance when it is in compliance. Not a good thing but its twin disagreement is worse where with false negatives it is decided that a rule/regulation is in compliance when it is out of compliance. False negatives need to be avoided because they place clients at extreme risk, more so than a false positive. False positives should also be avoided but it is more important to deal with the false negatives first before addressing the false positives.

Let’s look at this from a mathematical point of view in the following matrix. In order to better understand the above relationships and determine when ameliorative action needs to occur to shore up the differences between the agreements and disagreements, it is easier to do this mathematically than trying to eyeball it.

Table 2: Uncertainty-Certainty Matrix (UCM) Math Model

UCM Matrix Math Model		**Decision (D) Regarding**	**Regulatory Compliance**	**Totals**
		(+) In Compliance	(-) Not In Compliance	
Actual State (S)	(+) In Compliance	A	B	Y
Of Compliance	(-) Not In Compliance	C	D	Z
Totals		W	X	

Formulae based upon above: Agreements = (A)(D); Disagreements = (B)(C); Randomness = sqrt ((W)(X)(Y)(Z))

UCM Coefficient = ((A)(D)) - ((B)(C)) / sqrt ((W)(X)(Y)(Z)) in which a coefficient closer to 1 indicates agreement (certainty) and a coefficient closer to –1 indicates disagreement (uncertainty). A coefficient closer to 0 indicates randomness. Obviously, we want to see (A)(D) being predominant and very little in (B)(C) which are false positives and negatives where decisions and the actual state of regulatory compliance are not matching. If (WXYZ) is predominant then there is just randomness in the data. Also, not an intended result.

The reason for even suggesting this matrix is the high level of dissatisfaction with the levels of reliability in the results of program monitoring reviews as suggested earlier. If it were not so high, it would not be an issue; but with it being so high the field of licensing needs to take a proactive role in determining the best possible way to deal with increasing inter-rater reliability among licensing inspectors. Hopefully, this organizational schema via the UCM Matrix will help to think through this process related to licensing measurement and monitoring systems.

$$\boldsymbol{UCM} = \ll \boldsymbol{A} \times \boldsymbol{D} \gg - \ll \boldsymbol{B} \times \boldsymbol{C} \gg \div \sqrt{\_\_} \ll \boldsymbol{W} \times \boldsymbol{X} \times \boldsymbol{Y} \times \boldsymbol{Z} \gg$$

The above formula provides a means to calculate when action needs to be taken based upon the respective UCM coefficients. A UCM coefficient from +.25 to +1.00 is in the acceptable range; +.24 to -.24 is due to randomness and needs to be addressed with additional inter-rater reliability training; -.25 to –1.00 indicates a severe disagreement problem that needs to be addressed both in reliability training and a full review of the targeted rules/regulations to determine if the specific rule needs additional clarification.

Table 3: Uncertainty-Certainty Matrix (UCM) Licensing Decision Coefficient Ranges

UCM Coefficient	**Licensing Decision**
+.25 to +1.00	Acceptable, No Action Needed, In or Out of Regulatory Compliance Verified through mostly Agreements. (Generally, 90% of cases)
+.24 to -.24	Random, Agreements + Disagreements, Needs Reliability Training. (Generally, 5% of cases)
-.25 to –1.00	Unacceptable, Mostly Disagreements, Needs Training & Rule/Regulation Revision. (Generally, 5% of cases)

Figure 1: Kappa Coefficient

Observed agreement

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

Expected agreement if random judgment

Figure 2: Uncertainty-Certainty Coefficient

$$\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

Let’s provide an example of how this could work. A standard/rule/regulation that is common is the following:

Do all caregivers/teachers and children wash their hands often, especially before eating and after using the bathroom or changing diapers?

This is obviously an observation item where the licensing staff would observe in a sample of classrooms in a child care center for a set period of time. During their observations, there were several opportunities where the necessary behavior was required, and the staff complied with the rule and washed their hands. So, on the surface this specific rule was in compliance and there would appear to be full compliance with this rule based upon the observation.

A second scenario is where the observation is made, and the licensing staff observes the child care staff not washing their hands on several occasions. Then this specific rule would be out of compliance, and it would be duly noted by the licensing staff. These two scenarios establish a certain level of certainty during this observation session. However, there are other outcomes, for example, possibly one of the classrooms that was not observed had the opposite finding than what was observed in these particular classrooms. If data were being aggregated and a specific percentage was to be used the final decision about this rule could be different. Now we are getting into the uncertainty cells of the matrix where a false positive or negative could be the result. The licensing staff records the rule as being in compliance when in reality it is not = false negative or the rule is recorded as being out of compliance when in reality it is in compliance = false positive.

Another example which involves either Random Clinical Trials (RCT) or the use of abbreviated inspections (AI) and the results from these two interventions. The decision making in both RCT and AI is basically the same. We want to make sure that the results match reality. Every time an abbreviated review is done the following four regulatory compliance results should occur based upon the UCM matrix: 1) no additional random non-compliance is found; 2) there are no false negatives (abbreviated review finds no non-compliance but in reality there is); 3) when there is non-compliance found in abbreviated inspections, other related non-compliance is found; and 4) lastly the level of false positives (abbreviated review finds non-compliance but in reality there are no other related non-compliances) is kept to a minimum. This last result based upon copious research is that it is difficult to obtain but as the regulatory science moves forward hopefully this will become more manageable.

Hopefully these above examples provided some context for how the Uncertainty-Certainty Matrix (UCM) can be used in making specific licensing decisions based upon the regulatory compliance results.

1. Uncertainty-Certainty Matrix for Validation and Reliability Studies

The purpose of this part of this research abstract is to explore the possibility of utilizing the Uncertainty-Certainty Matrix (UCM) in validation and reliability studies in licensing decision making. The UCM has been proposed for use in licensing decision making but this would be an extension of this thinking to studies that involve validating licensing decisions such as when key indicators are used in comparison with comprehensive reviews of rules, and in reliability studies to determine individual inspector bias in regulatory compliance.

The basic premise of the UCM is that individual decision-making matches reality. When it comes to regulatory compliance decision making a 2 x 2 matrix can be drawn with the possible outcomes as is indicated in the following table (Table 4).

Table 4

UCM Matrix Logic		**Decision Regarding**	**Regulatory Compliance**
		(+) In Compliance	(-) Not In Compliance
Actual State of	(+) In Compliance	Agreement (++)	Disagreement (+-)
Compliance	(-) Not In Compliance	Disagreement (-+)	Agreement (--)

In using this table, the hope is that the decision regarding regulatory compliance matches the actual state of compliance where the coefficient is as close to +1.0 as possible, in other words, perfect agreement. So, the agreement cells are heavily weighted. We do not want to see all the cells, both agreement and disagreement cells, equally weighted. That would indicate a random response rate and a coefficient close to 0.0.

But there is another possibility which involves bias on the part of the licensing inspector in which they have certain biases or tendencies when it comes to making regulatory compliance decisions about individual rules. So, it is possible that decisions made regarding regulatory compliance could be either overall (+) positive In-Compliance or (-) negative Not-In-Compliance when in reality the actual state of compliance is more random.

When this occurs, the coefficient falls off the range category and is not between 0 and +/-1.0 because there is no variance detected in the data. It is always biased either positively or negatively.

The UCM can be used for both reliability and validity testing as suggested in the above. Just look for different results. For validity, false positives and negatives should either be eliminated or reduced as well as possible and the remaining results should show the typical diagonal pattern as indicated by the agreement cells.

For reliability, the same pattern should be observed as in the validity testing above but there is an additional test in which bias is tested for. Bias will be ascertained if the patterns in the results indicate a horizontal or vertical pattern in the data with little or no diagonal indication. Bias can be found at the individual inspector level as well as at the standard level or the actual state of compliance.

In both reliability and validity testing, random results in which each of the cells are equally filled is not a desirable result either.

The following tables 5-10 depict the above relationships with results highlighted in red:

Table 5

Valid & Reliable Results	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 6

Random Results	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 7

Positive Bias Results Individual	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 8

Negative Bias Results Individual	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 9

Positive Bias Results Standard	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 10

Negative Bias Results Standard	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Tables 5 – 10 demonstrate the different results based upon individual response rates when making regulatory compliance decisions about rules. Table 5 is what needs to be attained and tables 6 – 10 need to be avoided. Only in table 5 are false negatives and positives eliminated or avoided. In tables 6 – 10, false negatives and/or false positives are introduced which is not desirable when making validity or reliability decisions.

Table 6 results clearly indicate that a great deal of randomness has been introduced in the regulatory compliance decision making in which the individual licensing inspector decisions do not match reality. Tables 7 and 8, demonstrate bias in the decision-making process either positively (inspector always indicates in compliance) or negatively (inspector always indicates out of compliance). It is also possible that the standard being used has bias built into it, this is less likely but is still a possibility. The results in Tables 9 and 10 demonstrate where this could happen.

All these scenarios need to be avoided and should be monitored by agency staff to determine if there are patterns in how facilities are being monitored.

2. Uncertainty-Certainty Matrix for Differential Monitoring Studies

The purpose of this part of the research abstract is to explore the possibility of utilizing the Uncertainty-Certainty Matrix (UCM) not only in validation and reliability studies in licensing decision making but also with differential monitoring studies. The UCM has been proposed for use in licensing decision making but this would be an extension of this thinking to studies that involve validating licensing decisions such as when key indicators are used in comparison with comprehensive reviews of rules, and in the development of risk rules as part of the risk assessment methodology. This new Differential Monitoring 2x2 Matrix can also be used to depict the relationship between full and substantial regulatory compliance and the nature of rulemaking.

The basic premise of the DMM: Differential Monitoring Matrix is similar to the original thinking with the UCM but there are some changes in the formatting of the various cells in the matrix (see Table 11). When it comes to regulatory compliance decision making a 2 x 2 matrix can be drawn with the possible outcomes as is indicated in Table 11 where each individual rule is either in (+) or out (-) of compliance. Also, there is the introduction of a high regulatory compliant group (+) and a low regulatory compliant group (-) which is different from the original UCM.

Table 11

DMM Matrix	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

By utilizing the format of Table 11, several key components of differential monitoring can be highlighted, such as key indicators and risk assessment rules, as well as the relationship between full and substantial regulatory compliance.

Regulatory compliance is grouped into a high group (+), generally this means that there is either full or substantial regulatory compliance with all rules. The low group (-) usually has 10 or more regulatory compliance violations. Individual rules being in (+) or out (-) of regulatory compliance is self-explanatory.

Tables 12-18 below will demonstrate the following relationships:

Table 12 depicts the key indicator relationship between individual rules and the high/low groups as indicated in red. In this table, the individual rule is in compliance with the high group and is out of compliance with the low group. This result occurs on a very general basis and should have a .50 coefficient or higher with a p value of less than .0001.

Table 13 depicts what most rules look like in the 2x2 DMM. Most rules are always in full compliance since they are standards for basic health and safety for individuals. This is especially the case with rules that have been weighted as high-risk rules. Generally, one never sees non-compliance with these rules. There will be a substantial number of false positives (+-) found with high-risk rules but that is a good thing.

Table 14 depicts what happens when full compliance is used as the only criterion for the high group. Notice that the cell right below (++) is eliminated (-+). This is highly recommended since it eliminates false negatives (-+) from occurring in the high group. As will be seen in Table 5, when substantial compliance is used as part of the high group sorting, false negatives are re-introduced. If possible, this should be avoided, however in some cases because of the regulatory compliance data distribution it is not always possible where not enough full compliant programs are present.

Table 15 depicts what occurs when substantial compliance is used as part of determining the high group. False negatives can be reintroduced into the matrix which needs to be either eliminated or reduced as best as possible. If substantial compliance needs to be used in determining the high group, then there is a mathematical adjustment that can be made which will impact the equation and essentially eliminate false negatives mathematically (see the research note at the end of this research abstract).

Table 16 depicts what happens if the individual rule is particularly difficult to comply with. Both the high performers as well as the low performers are out of compliance with the rule.

Table 17 depicts a situation where the programs are predominantly in a low group with few at full or substantial regulatory compliance which is indicative of poor performing programs. Very honestly, this is generally not seen in the research literature, but it is a possibility and one to be in tune with.

Table 18 depicts a terrible individual rule which predicts just the opposite of what we are trying to do with programs. Obviously, this rule would need to be rewritten so that it fits with the essence of regulatory compliance in helping to protect individuals.

The following tables 12-18 will depict the above relationships with results highlighted in red:

Table 12

Key Indicators	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 13

Risk Rules	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 14

Full Compliance	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance		(--)

Table 15

Substantial Compliance	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 16

Very Difficult Rule	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 17

Poor Performing Programs	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 18

Terrible Rule	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Tables 12 – 18 demonstrate the different results based on the relationship between individual regulatory compliance and if a program is either a high performer or a low performer. These tables are provided as guidance for understanding the essence of differential monitoring and regulatory compliance which has various nuances when it comes to data distributions. This research abstract hopefully can be used as a guide in determining from a data utilization point of view how to make important regulatory compliance policy decisions, such as: which rules are excellent key indicator rules, which are performing as high risk rules, importance of full compliance, what to do when substantial compliance needs to be employed, are there difficult rules to comply with, how well are our programs performing, and do we have less than optimal rules that are in need of revision.

Research Note:

Over the past decade in doing research on the Regulatory Compliance Key Indicator Metric (RCKIm) it has become very clear that false negatives needed to be controlled for because of their potential to increase morbidity and mortality. When dealing with regulatory compliance and full compliance as the threshold for the high grouping variable in the 2 x 2 Regulatory Compliance Key Indicator Matrix (RCKIm) (see matrix below in Table 19), false negatives could be either eliminated or reduced to the point of no concern.

However, if substantial compliance rather than full compliance is used as the threshold for the high grouping variable in the 2 x 2 Regulatory Compliance Key Indicator Matrix (RCKIm) this becomes a problem again. There is the need to introduce a weighting factor. In utilizing the RCKIm, the following equation/algorithm is used to produce the Fiene Coefficient (FC):

FC = ((A)(D)) – ((B)(C)) / sqrt (WXYZ)

This RCKIm needs to be revised/updated to the following to consider the need to again eliminate false negatives being generated by the results of the equation/algorithm; this can be accomplished by cubing B:

FC* = ((A)(D)) – ((B^3)(C)) / sqrt (WXYZ)

By this simple adjustment to cube (B = False Negatives) it will basically eliminate the use of any results in which a false negative occurs when substantial compliance is determined. The table below (Table 19) displays the variables of the Regulatory Compliance Key Indicator Matrix (RCKIm).

Table 19: RCKIm	High RC Group	RC Low Group	
KI In Compliance	A	B^3	Y
KI Violations	C	D	Z
Totals	W	X	

In the above examples, FC can be used when the High RC Group is at full regulatory compliance, but FC* needs to be used when the High RC Group is including substantial as well as full regulatory compliance. By using both equations/algorithms, it better deals with the results of the Regulatory Compliance Theory of Diminishing Returns.

The results should clearly show that only positive (+) coefficients will become Regulatory Compliance Key Indicators versus those rules that do not show any relationship to overall regulatory compliance (0), but now the negative (-) coefficients will more clearly show when any false negatives appear and clearly not include them as Regulatory Compliance Key Indicators. This is a major improvement in the Regulatory Compliance Key Indicator methodology which clearly demonstrates the differences in the results. It provides a gateway in regulatory compliance data distributions where substantial regulatory compliance is heavily present while full regulatory compliance is not. This could become a problem as the regulatory science field moves forward with the use of the Regulatory Compliance Theory of Diminishing Returns.

3. Uncertainty-Certainty Matrix Logic Model and Algorithms

The Uncertainty-Certainty Matrix (UCM) has been introduced for licensing decision making and validating differential monitoring approaches. This section builds upon this introduction by presenting a logic model and its corresponding algorithms. The purpose of this section is to expand that original thinking and to introduce the associated algorithms that specify the UCMLMA (Uncertainty-Certainty Matrix Logic Model and Algorithms).

The UCMLMA will build off the contingency table and confusion matrix logic modeling as described above (in tables 1 and 2) in the following matrix (Table 20).

Table 20: UCMLMA: Uncertainty-Certainty Matrix Logic Model and Algorithms

UCMLMA	Actual:	Individual	Inspector
Reality:	Decisions:	C = Compliance	NC = Non Compliance
Gold	C = Compliance	*TP = True Positive*	*FP = False Positive*
Standard	NC = Non Compliance	*FN = False Negative*	*TN = True Negative*

From this matrix certain licensing decision making can be made and certain biases can be avoided with the following algorithms related to compliance and non compliance of rules/regulations.

TP + TN is the ideal situation in which all decisions are made correctly. Accuracy can be measured by *((TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* with a coefficient closer to +1.00.

TP + FP + FN + TN is a totally random situation where decisions are not reliable nor valid, there is no rhyme or reason to the decision making process by the individual inspector. Randomness can be measured by the following: *(TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* with a coefficient closer to 0.00.

TP + FP introduces a positive bias in which the inspector has a tendency to always make a decision in which the program is in compliance with specific rules. They are overly lenient in their interpretation of the rules. Positive bias sensitivity can be measured by *FP / (TP + FP).* The higher the percent, the more bias is present.

FN + TN introduces a negative bias in which the inspector has a tendency to always make a decision in which the program is in non compliance with specific rules. They are overly stringent in their interpretation of the rules. Negative bias sensitivity can be measured by *FN / (FN + TN).* The higher the percent, the more bias is present.

Table 21: UCMLMA: Uncertainty-Certainty Matrix Logic Model and Algorithms Applied to Differential Monitoring

UCMLMA		Overall	Compliance
Individual	Decisions:	High Group (top 10%)	Low Group (bottom 10%)
Rule/Standard/	C = Compliance	*TP = True Positive*	*FP = False Positive*
Regulation	NC = Non Compliance	*FN = False Negative*	*TN = True Negative*

From this matrix (Table 21), certain decisions can be made regarding differential monitoring validation related to key indicators, risk assessment, regulatory compliance levels and individual rule performance with the following algorithms.

TP + TN is the ideal result when determining key indicators because the individual rule statistically predicts overall compliance with all rules. Accuracy can be measured with the following: *((TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* with the coefficient being closer to +1.00.

TP + FP is what happens with high risk rules in that they are always in compliance and one sees very little non compliance with these high risk rules generally. Sensitivity can be measured by *FP / (TP + FP)*. Higher compliance rates would generally indicate higher risk rules.

TP is when 100% full compliance is always present. The greater the number or percent of total programs being 100% in full compliance, the greater the skewness in the data distribution.

TP + FN is present when substantial regulatory compliance is used as a criterion for licensing decision making. The higher the number, the greater the skewness in the data distribution.

FN + TN occurs when a rule is difficult to comply with. The following algorithm can be used *FN / (FN + TN)* to measure how difficult the rule is. The greater the percent, the more difficult the rule is.

FP + TN occurs when programs are performing poorly and there is a great deal of non compliance with the rules. The greater the number of programs indicates a poorly performing system of rules.

FP + FN is an example of a terrible rule that predicts the opposite of what we intended with overall regulatory compliance. The following algorithm *((TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* can be used with coefficients closer to -1.00 demonstrating this finding.

The above two tables (table 20 and 21) and their corresponding algorithms give a structured approach to licensing decision making and differential monitoring validation. It should provide the licensing administrator with a data driven and empirical method based upon regulatory science principles. By utilizing the concepts, logic model, and algorithms presented in this working paper, it will hopefully add to a regulatory scientific approach in the human services licensing and regulatory administration field when it comes to making licensing decisions that avoid bias and reliability & validity errors in decision making.

How the Theory of Regulatory Compliance Explains the Relationship Between Structural and Process Quality, Data Distributions, Scoring Systems, and a New Scale for Parents

Richard Fiene PhD

Penn State Edna Bennett Pierce Prevention Research Center

June 2025

Child care and early education (CCEE) quality has been defined in the research literature along a structural and process continuum where structural quality has been dealing with hard, countable standards while process quality deals with the softer side of quality dealing with adult child interactions. To add more substance to this continuum, process quality is the real heart of quality, getting at the essence of those intricacies of what happens in individual classrooms in individual and group interactions amongst teachers and children. Structural quality are the surrogates to quality, such as measuring compliance with the number of teachers to children in staff-child ratios, or group sizes of children, or the number of violations of specific rules, regulations or standards. Structural quality does not look at the softer elements of quality when it comes to interactions or classroom atmosphere, sometimes it looks at the program curriculum but generally not. Structural quality is more concerned with health and safety standards, things that may harm children rather than things that will enhance their environment, that is left to process quality.

Structural quality elements are generally present in licensing rules and regulations, while process quality elements are present in tools such as the Environmental Rating Scales (ERS) or the Classroom Assessment Scoring System (CLASS). The ERS and CLASS are generally not used on their own although that was their original intent but are usually a part of other quality initiatives or CCEE systems, such as QRIS (ERS) and Head Start (CLASS). Structural and process quality complement each other in a building block way. Structural quality provides the foundation while process quality builds upon that foundation in an ever-expanding manner.

Another way of thinking about quality and its elements, is to think of a quality spectrum where we place structural and process quality on a spectrum line with the associated quality interventions. The quality interventions can be grouped in the following manner for structural quality: they would include licensing, quality rating and improvement systems (QRIS), Head Start Performance Standards, accreditation, and professional development systems. For process quality this is where the ERS and CLASS tools would go. Think of the quality spectrum as using a prism and splitting up light into all its various wavelengths and resulting colors.

How does the theory of regulatory compliance fit into all this? The theory provides the overarching and unifying framework to depict how structural quality and process quality work together. One of the main discoveries with the theory of regulatory compliance was demonstrating the importance of substantial regulatory compliance with structural quality rules. This discovery was made when a ceiling effect was determined in comparing structural to process quality. And this ceiling effect was discovered in all structural quality: licensing, Head Start, accreditation, and QRIS systems. Licensing demonstrates the greatest ceiling effect and, in some cases, a diminishing returns effect when moving from substantial to full 100% compliance but all of these structural quality systems demonstrate some form of a ceiling effect. Process quality follows a linear relationship, and its data distribution is normally distributed while structural quality follows a nonlinear relationship, and its data distribution is positively skewed. Studies in CCEE over the past 50 years have clearly demonstrated these relationships with structural and process quality when it comes to measuring compliance with the rules, regulations, and standards of each view of quality.

The theory of regulatory compliance has led to refocusing licensing decision-making that takes substantial compliance into account when determining who gets a full license and who does not. It clearly demonstrates how at times substantial compliance is equivalent to full 100% compliance with all rules, regulations, or standards; and, in some cases, is better than full compliance. This has also led to abbreviated, targeted or focused inspections where key predictor rules or high-risk rules are assessed which instituted a nuanced program monitoring system called differential monitoring.

It has also led to identifying quality indicators and infusing quality into the licensing rule and regulatory landscape. The use of licensing and quality predictor indicators has been the cornerstone of the differential monitoring approach and for good reason. These licensing and quality indicators can be looked upon as the anchors to structural and process quality. These key indicators statistically predict overall compliance with the full set of rules, regulations, and standards and studies have confirmed this relationship in licensing repeatedly, QRIS, Head Start, accreditation, ERS, and in the development of a new quality indicator scale.

From a statistical methodological point of view, it has led to, at times, significant correlations between structural and process quality but generally these correlations are at the lower end of significance. The reason being is that structural quality follows this ceiling effect or nonlinear skewed data distribution which does not match with the normal distributions found in process quality data distributions. So, researchers and scientists should not be surprised to find that their correlations between process and structural quality elements are not statistically significant. When looking at structural quality it is difficult to distinguish between the truly high performers and the mediocre performers. With process quality, it is much easier making that determination. With both structural and process quality it is equally easy to distinguish the high performers from the low performers.

Because of this difficulty in distinguishing between the high performers and the mediocre performers has led to the introduction of a new metric in structural quality called the Regulatory Compliance Scale (RCS). The reason for doing this is twofold: 1) The RCS fits more closely with the theory of regulatory compliance in demonstrating the importance of substantial compliance and having a categorical sequencing; 2) The categorical or ordinal sequence fits nicely with the existing process quality tools which are organized and measured on an ordinal scale of 1 - 7 scale. The RCS has been pilot tested in several jurisdictions, and it has demonstrated its ability to be a better measure when comparing structural quality to process quality than using straight rule, regulation or standard compliance violation frequency data.

These above assertions have been addressed previously but probably not in one place demonstrating the impact of the theory of regulatory compliance on structural and process quality. It is hopeful in the coming years that research psychologists and regulatory scientists will attempt to replicate these findings so that the public policy implications can be carried to their logical end point: substantial compliance being a sufficient level of compliance for issuing a full license, and the institutionalization of differential monitoring throughout the CCEE field. For this to happen, the ceiling effect in structural quality needs to be replicated when compared to process quality.

1. Structural Quality (RC) and Process Quality (PQ) Data Distributions

This research abstract provides the data distributions for a series of structural (RC) and program quality (PQ) studies which show dramatically different frequencies and centralized statistics. The structural quality data distributions have some very important limitations that will be noted as well as some potential adjustments that can be made to the data sets to make statistical analyses more meaningful. These data distributions are from the USA and Canada.

It is obvious when one observes the PQ as versus the RC data distributions that the RC data distributions are much more skewed, medians and means are significantly different, and kurtosis values are much higher which means that the data contain several outliers. These data distributions are provided for researchers who may be assessing structural quality (RC) data for the first time. There are certain limitations of these data which are not present in more parametric data distributions which are more characteristic of process quality (PQ) data.

To deal with the level of skewness of RC data, weighted risk assessments have been suggested in order to introduce additional variance into the data distributions. Also, dichotomization of data has been used successfully with very skewed data distributions as well. One of the problems with very skewed data distributions is that it is very difficult to distinguish between high performing providers and mediocre performing providers because of the ever present ceiling effect when comparing structural quality to process quality. Skewed data distributions provide no limitations in distinguishing low performing providers from their more successful providers.

In looking at the process quality (PQ) data, these data are generally normally distributed and do not demonstrate any severe skewness in their respective distributions. These data are distributed as should be the case where sufficient variance is present and there is no need for weighting because of the dispersion in the data. It is easy to be able to distinguish high performing providers from mediocre performing providers and from low performing providers.

For purposes of reading the following Table:
Data Set = the study that the data are drawn from.
Sites = the number of sites in the particular study.
Mean = the average of the scores.
sd = standard deviation.
p0 = the average score at the 0 percentile.
p25 = the average score at the 25th percentile.
p50 = the average score at the 50th percentile or the median.
p75 = the average score at the 75th percentile.
p100 = the average score at the 100th percentile.

Data Set	Sites	mean	sd	p0	p25	p50	p75	p100	PQ or RC
ECERS total score PQ	209	4.24	0.94	1.86	3.52	4.27	4.98	6.29	PQ
FDCRS total score PQ	163	3.97	0.86	1.71	3.36	4.03	4.62	5.54	PQ
ECERS and FDCRS totals PQ	372	4.12	0.91	1.71	3.43	4.12	4.79	6.29	PQ
ECERS prek PQ	48	4.15	0.74	2.56	3.6	4.15	4.65	5.56	PQ
ECERS preschool PQ	102	3.42	0.86	1.86	2.82	3.26	4.02	5.97	PQ
ITERS PQ	91	2.72	1.14	1.27	1.87	2.34	3.19	5.97	PQ
FDCRS PQ	146	2.49	0.8	1.21	1.87	2.42	2.93	4.58	PQ
CCC RC	104	5.51	5.26	0	2	4	8	25	RC
FCC RC	147	5.85	5.71	0	2	4	8.5	33	RC
CCC RC	482	7.44	6.78	0	2	6	11	38	RC
FDC RC	500	3.52	4.05	0	0	2	5	34	RC
CI Total Violations RC	422	3.33	3.77	0	1	2	5	24	RC
CLASS ES PQ	384	5.89	0.36	4.38	5.69	5.91	6.12	6.91	PQ
CLASS CO PQ	384	5.45	0.49	3.07	5.18	5.48	5.77	6.56	PQ
CLASS IS PQ	384	2.98	0.7	1.12	2.5	2.95	3.37	5.74	PQ
CLASS TOTAL OF THREE SCALES PQ	384	14.33	1.32	8.87	13.52	14.33	15.11	17.99	PQ
ECERS PQ	362	4.52	1.05	1.49	3.95	4.58	5.25	7	PQ
FDCRS PQ	207	4.5	1	1.86	3.83	4.66	5.31	6.71	PQ
CCC RC	585	5.3	5.33	0	2	4	8	51	RC

Studies Completed After 2020:

QRIS RC	585	2.78	1.24	0	2	3	4	4	RC
FDC RC	2486	2.27	3.42	0	0	1	3	34	RC
FDC PQ	2486	1.35	1.26	0	0	1	2	4	PQ
CCC RC	199	7.77	8.62	0	3	6	10	61	RC
CCC RC	199	6.69	10.32	0	1	4	8	98	RC
CCC RC	199	6.77	7.91	0	1.5	4	8.5	57	RC
QRIS RC	199	1.06	1.32	0	0	1	2	4	RC
CCC RC	199	7.08	6.96	0	2.33	5.67	9.84	52	RC
QRIS RC	381	2.55	0.93	0	2	3	3	4	RC
CCC RC	1399	1.13	2.1	0	0	0	1	20	RC
CCC RC	153	5.28	5.97	0	1	3	6	32	RC
FDC RC	82	3.52	4.36	0	0	2	4	21	RC

2. Structural and Process Quality Scoring Systems in Child Care and Early Education

This section will delve into the details of the scoring systems used in structural and process quality for child care and early education programs. The scoring systems have evolved significantly over the years influenced by the tests and measurement research literature. Presently a significant change has been proposed with measuring structural quality that needs to be shared and reacted upon.

There are a great deal of similarities between the scoring systems for structural quality and process quality although the content varies greatly. Both structural and process quality scoring systems measure compliance with specific rules, items, or standards in a similar fashion. Both can use a weighting of each rule, item or standard but that is not always the case. And usually the overall score or global score is on a scale such as 1-7 (Environmental Rating Scales (ERS) and Classroom Assessment Scoring System (CLASS)) or 1-5 (Quality Rating and Improvement Systems (QRIS)) with the exception of licensing systems in which violation counts have been used in the past.

The purpose of this section is to suggest the use of a scale (Regulatory Compliance Scale (RCS)) in place of doing violation counts so that licensing data distributions can mirror more closely what is occurring in other structural quality systems, such as QRIS and accreditation systems; and throughout process quality systems (ERS and CLASS). The reason for suggesting this change is that in studies conducted in the state of Washington and the Province of Saskatchewan it was determined that the RCS was more effective in distinguishing the relative quality of programs than using violation count data.

Here is a potential scale structure that could be used in transposing the violation count data to the Regulatory Compliance Scale (RCS) (Table 1). These thresholds or buckets for the RCS scale were determined by analyzing a multitude of regulatory compliance data sets drawn from the USA and Canadian Provinces and these violation counts were found to be best at distinguishing the various levels of quality.

Table 1: Comparison of RCS with Regulatory Compliance Violation Counts

RCS	Violation Counts	Description
7	0	Full 100% Regulatory Compliance
5	1-2	Substantial Regulatory Compliance
3	3-9	Mediocre Regulatory Compliance
1	10+	Low Non-Optimal Regulatory Compliance

This transformation of data from violation counts to the RCS scale could vary if a weighting system is used with the rules. If not, then the above transformation has worked well in creating this ordinal/categorical ranking of regulatory compliance related to violation count data that has not been weighted. It fits with the prevailing theory of regulatory compliance which discusses the importance of substantial compliance with rules in determining the quality of a setting.

By utilizing this change in moving to a Regulatory Compliance Scale, it mirrors the other scoring systems in structural quality and all the scoring systems in process quality. From an analytical point of view, this greatly simplifies and makes more straightforward future analyses.

3. The Emergence of a New Early Childhood Program Quality Tool/Scale for Parents in Measuring both Structural and Process Quality in Selecting Child Care

This section provides an overview that parents can look for in their selection process of choosing high quality child care for their children. It is based upon 50 years of research delving into what constitutes a high quality child care and early education program. It is drawn from the major quality initiatives that have been implemented throughout the USA and Canada during this time frame.

The key indicators that every parent should be looking for:

1) You will always start with the staff, the early childhood educators and determine if they have the necessary credentials in early childhood education. The best indicators are teachers having CDA, AA or BA degrees in early childhood education.

2) Look to see if there is a program curriculum and if that program is child centered. Children should be viewed as competent learners, and they have the freedom to access classroom materials independently without adult intervention. The children are provided with meaningful choices through activity/learning centers. There is evidence of the children's interests and their projects in the learning environment.

3) Does the program follow an individualized prescribed planning document when it comes to curriculum. It does not mean it is a canned program, in fact, it shouldn't if it is based upon the individual needs of each child's developmental assessment. There should be a written document that clearly delineates the parameters of the philosophy, activities, guidance, and resources needed for the particular curricular approach. There should also be a developmental assessment which is clearly tied to the curriculum. The developmental assessment can be home-grown or a more standardized off-the- shelf type of assessment, the key being its ability to inform the various aspects of the curriculum. The purpose of the assessments is not to compare children but rather to compare the developmental progress of individual children as they experience the activities of the curriculum. The program practices emergent curriculum, allowing the interests of the children to determine the learning content. The curriculum is informed by individual developmental assessments of each child in the respective classrooms. The children and educators are co-learners in the exploration of projects. Learning activities of the children are documented, displayed in the learning environment and used to plan further learning activities. This can be assessed developmentally.

4) There should be activities both within the center as well as off site where teachers and parents have opportunities to meet and greet each other. Communication with family members is documented and enables early childhood providers to assess the need for follow-up. Early childhood providers hold regular office hours when they are available to

talk with family members either in person or by phone. Family members are encouraged to lead the conversation and to raise any questions or concerns.

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

6) Educators Encourage Children to Communicate. Conversations and questions should be used with all children, even young infants. Conversations using verbal and nonverbal turn-taking should be considered when scoring. Most conversations and questions initiated by infants will be nonverbal, such as widening of the baby's eyes or waving arms and legs. Observe staff response to such nonverbal communication. For infants and toddlers, the responsibility for starting most conversations and asking questions belongs to the staff. As children become more able to initiate communication, staff should modify their approach in order to allow children to take on a greater role in initiating conversations and asking questions. Staff should provide answers to questions used by children if children cannot answer, and as children become more able to respond, questions should start to include those that the child can answer.

7) Educators Use Language to Develop Reasoning Skills. Staff should use language to talk about logical relationships using materials that stimulate reasoning. Through the use of materials, staff can demonstrate concepts such as same/different, classifying, sequencing, one-to-one correspondence, spatial relationships, and cause and effect.

8) Educators Listen Attentively When Children Speak. Do the educators look directly at the children with nods, rephrase their comments, engage in conversations. Children should have the undivided attention of the specific educator they are addressing. Staff should not be looking away or preoccupied with others. They should be at the child's level making eye contact.

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

These are the basic structural and process quality key indicators that every parent should be looking for when selecting their child care. The more of these you see the better.

Development of a Regulatory Compliance Scale

Richard Fiene PhD

Penn State's Edna Bennett Pierce Prevention Research Center

April 2025

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

So how would this look from a licensing decision making vantage point. Presently, in licensing measurement, licensing decisions are made at the rule level in which each rule is either in or out of compliance in the prevailing paradigm. Licensing summaries with corrective actions are generated from the regulatory compliance review. It is a nominal measurement system 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 both be measured by licensing inspectors, there would be no need to have two separate systems but rather one that applies to everyone and becomes mandated rather than voluntary. It could help to balance both effectiveness and efficiency by only including those standards and rules that statistically predict regulatory compliance and quality and balancing risk assessment by adding high risk rules.

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

Introduction to the Regulatory Compliance Scale

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

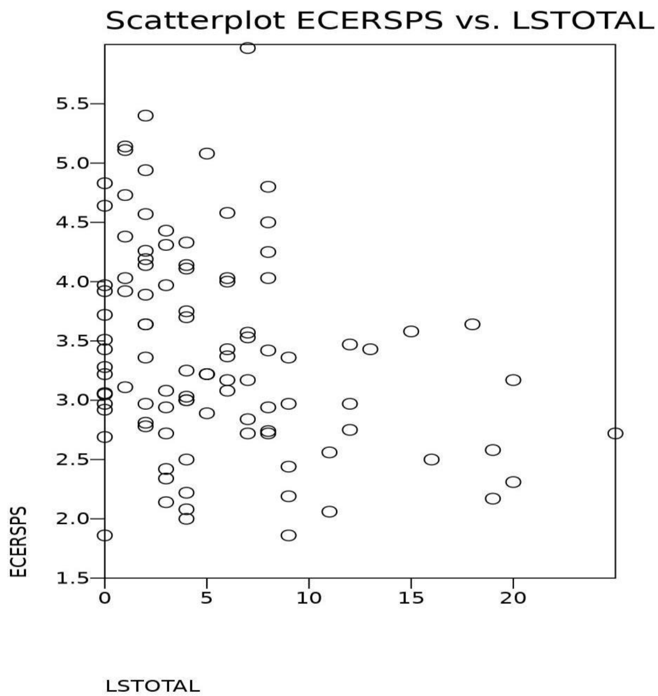
The theory of regulatory compliance has another application when it comes to regulatory compliance measurement in helping to move the licensing field from a nominal based measurement strategy to one of ordinal based measurement. The new measurement strategy is the Regulatory Compliance Scale (RCS) and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

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

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

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

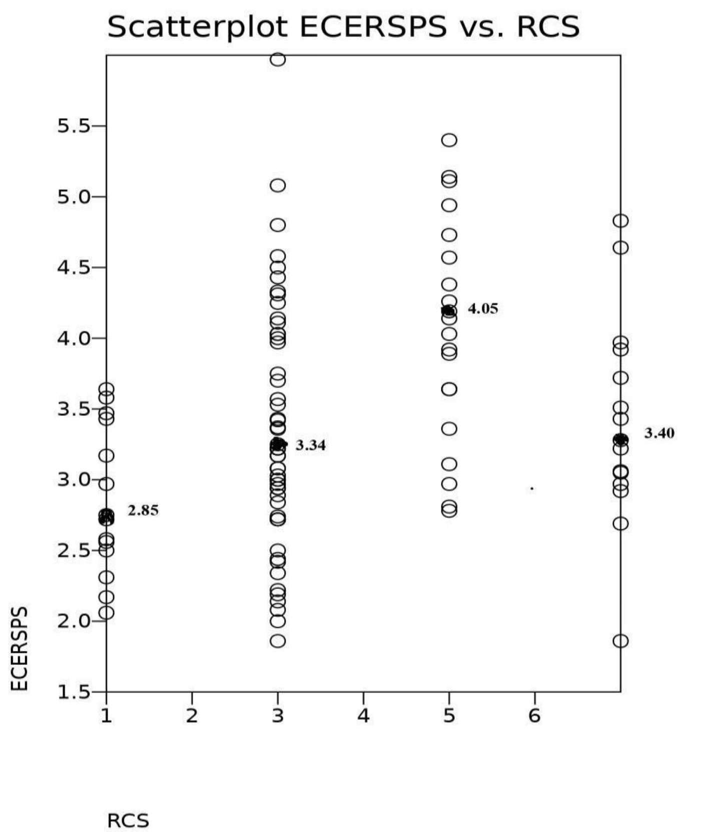
The first figure shows the actual individual violation data of the programs compared to their corresponding ECERS scores. There is not a significant relationship between the two as depicted in the graphic.



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

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

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



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

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

Regulatory Compliance Scale Studies

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

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

Methodology

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

RCS Models used for analyses

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

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

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

Results

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

RCS Model Results compared to Quality Scales

RCS results	**Models**	**Quality1**	**Quality2**	**Quality3**
Jurisdiction1	**RCS0**	**.26***	**.39***	**.39***
	RCS3	.21	.32*	.33*
	RCS5	.20	.36*	.33*
Jurisdiction2	**RCS0**	**.76****	**.46****	**---**
	RCS3	.12	-.07	---
	RCS5	.18	-.02	---
	RCSF1	.55**	.29*	---
	RCSF2	.63**	.34	---
Jurisdiction3	RCS0	.19	.18	.16
	RCS3	**.21**	**.21**	**.15**
	RCS5	.18	.16	.07
	RCSF1	.17	.17	.10
	RCSF2	.18	.18	.19
Jurisdiction4	RCS0	.24*	---	---
	RCS3	.28*	---	---
	RCS5	**.30***	**---**	**---**
	RCSF1	.21	---	---
	RCSF2	.29*	---	---
Jurisdiction5	RCS0	.06	-.02	.07
	RCS3	.06	-.01	.05
	RCS5	**.08**	**.00**	**.09**
	RCSF1	.00	-.03	.05
	RCSF2	.05	-.03	.05

*Statistically significant .05 level;

**Statistically significant .01 level.

In the above table starting under Jurisdiction2, two new models were introduced based upon the Fibonacci Sequence (Fibonacci1 = RCSF1; Fibonacci2 = RCSF2) and their model structure is in the following Table. The reason for doing this is that the Fibonacci Sequence introduces additional variation into the scaling process.

RCS Fibonacci Models

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

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

ANOVAs Comparing the RCS Models with Program Quality

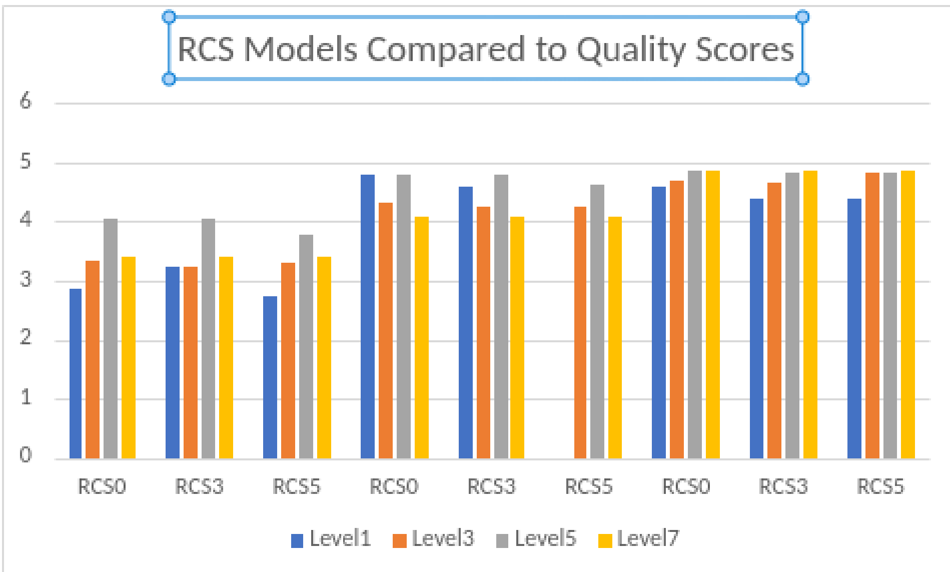
Jurisdictions	**Model**	**Level 1**	**Level 3**	**Level 5**	**Level 7**
Jurisdiction1	**RCS0**	**2.85**	**3.34**	**4.05**	**3.40**
	RCS3	3.24	3.23	4.05	3.40
	RCS5	2.73	3.32	3.77	3.40
Jurisdiction2	RCS0	4.81	4.31	4.80	4.10
	RCS3	**4.59**	**4.25**	**4.80**	**4.10**
	RCS5	---	4.26	4.64	4.10
Jurisdiction3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

Discussion

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

The bottom line is that the original RCS Model is as good as any and no other model is consistently better than all the rest. The RCS Model does have a slight edge over Regulatory Compliance Violation RCV frequency counts in some jurisdictions but not in others. It is much easier to interpret the relationship between quality and the RCS models than it is to interpret the results from the quality scores and the RCV data distribution. So, the recommendation would be for licensing agencies to think about using this new scaling technique in one of its model formats to determine its efficacy. Pairing up RCS and RCV data side by side by licensing agencies would be important studies to determine which approach is the better approach.

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



Additional Analyses Comparing the 11 Studies

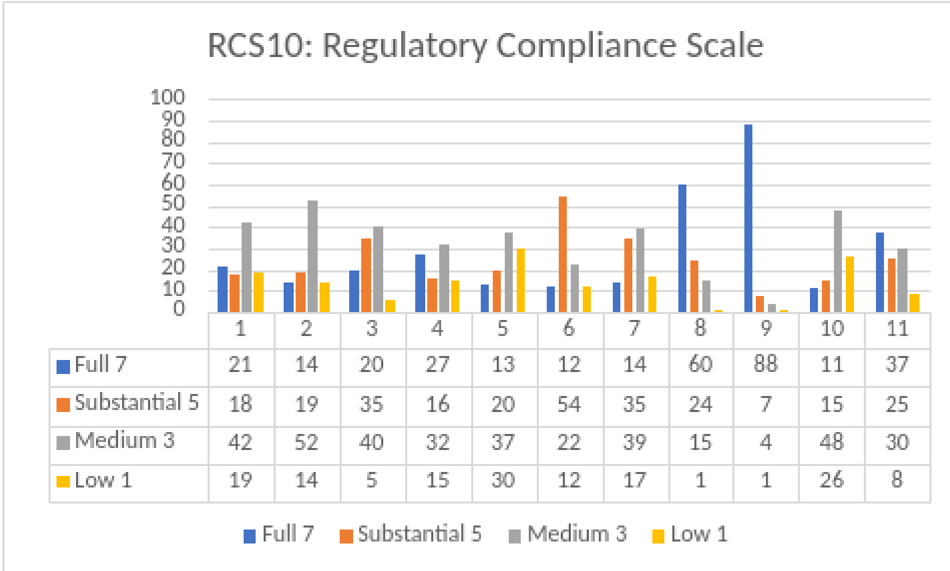
This section provides the results from 11 studies from 10 states and Canadian Provinces in which the proposed new Regulatory Compliance Scale (RCS) was utilized as a byproduct of a differential monitoring implementation or validation study. These studies were undertaken over a decade long period (2013-2023).

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

These are the results from these 10 jurisdictions which are presented in the following Table (all results are presented as percents of programs that fell into the scaling 1-7). Under the Studies, the number of the specific study is provided, followed by the sample size, followed by if it is in the USA (US) or Canada (CA).

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

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



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

Regulatory Compliance Scale Extensions

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

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

It is proposed in order to mitigate the plateau effect with regulatory compliance and program quality standards because regulatory compliance data distributions are severely skewed which means that many programs that have questionable quality are being included in the full (100%) compliance domain. When regulatory compliance standards are increased in their quality components this will lead to a higher level of overall quality as depicted in the "**XX**" cell all the way on the lower right. It also helps to mi gate the severe skewness in the regulatory compliance data distribution. The data distribution does not approximate a normally distributed curve which is the case with the program quality data distribution.

Regulatory Compliance x Program Quality Grid Model

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

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

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

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

Regulatory Compliance Scoring System and Scale (RC3S)

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

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

Regulatory Compliance Scaling for Decision Making

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

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

The proposed scaling has three major decision points attached to regulatory compliance scores. Either programs are in full or substantial compliance, in low compliance or somewhere in the middle. Full or substantial regulatory compliance is 100% or 99-98% in regulatory compliance. Low regulatory compliance is less than 90% and mid-regulatory compliance is between 97%-90%. These ranges may seem exceptionally ght but based upon the national database on regulatory compliance that I maintain at the Research Institute for Key Indicators (RIKILLC) these are the ranges that have formed over the past 40 years. These data ranges should not come as a surprise because we are talking about regulatory compliance with health and safety standards. These are not quality standards; these are basic protections for clients. The data are not normally distributed, not even close as is found in quality tools and standards.

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

Data	Level	Decision
100-98%	***Full/Substantial***	***License***
97-90%	***Mid-Range***	***Provisional***
89% or less	***Low***	***No-License***

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

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

Conclusion

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

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

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

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

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

Without the Theory of Regulatory Compliance, differential and integrative monitoring would not be needed because regulatory compliance would have had a linear relationship with program quality and full compliance would have been the ultimate goal. There would have been no need for targeted rule enforcement or reviews because all rules would have had an equal weight when it came to protecting clients and any individual rule would have predicted overall compliance. But it "just ain't so" as it is said. The need to make adjustments is brought about by the theory and it has not been the same ever since.

Regulatory Compliance Scales and Program Monitoring Systems

Scoring Level	**Individual Rule**		**Aggregate Rules**	**Individual Rule**
Scale	**Instrument based**	***Scale***	**Differential**	**Integrated**
7	Full Compliance	7	Full Compliance	Exceeds Compliance
-	---	5	Substantial	Full Compliance
-	---	3	Mediocre	Substantial
1	Out of Compliance	1	Low	Mediocre/Low

The above table attempts to summarize in tabular form the previous paragraphs in describing the relationship between program monitoring and licensing measurement scaling via a proposed regulatory compliance scale. As one can see this moves the paradigm from a nominal to an ordinal measurement rubric and depicts the differences in the measurement focus either at the individual rule or aggregate rules scoring levels. It also considers the significance of substantial compliance given the theory of regulatory compliance in which substantial compliance focus is a "*sweet spot*" phenomenon as identified in the regulatory science research literature. It is hoped that the regulatory science field takes these paradigm shifts into consideration in moving forward with building licensing decision making systems and how licenses are issued to facilities.

As a final footnote, keep in mind that the Theory of Regulatory Compliance applies to the relationship between regulatory compliance and program quality and does not apply to regulatory compliance in and of itself related to health and safety. When dealing with regulatory compliance, full compliance is the ultimate goal with individual rules and in determining which rules are predictive rules. It is the preferred methodology in order to eliminate false negatives and decreasing false positives in making licensing decisions related to regulatory compliance.

These above concepts all relate to the field of regulatory compliance and how to make informed decisions about licensing, particularly in the context of program monitoring. Here's how they connect:

Regulatory Compliance Scales:

These scales move away from a binary "compliant" or "non-compliant" approach to regulations. Instead, they acknowledge degrees of compliance, recognizing that minor deviations may not be as detrimental as major ones.

They provide a framework for evaluating the severity and frequency of non-compliance, allowing for more nuanced licensing decisions.

Instrument Based Program Monitoring (IBPM):

This is the traditional method of monitoring compliance, relying on standardized instruments and checklists to assess adherence to specific rules.

It's a comprehensive approach, but can be time-consuming and inflexible, potentially leading to over-regulation or missing important aspects of program quality.

Differential Monitoring (DM):

This approach takes into account the risk associated with different regulations, focusing monitoring efforts on areas with the highest potential for harm or non-compliance.

It allows for a more efficient use of resources and can be tailored to the specific needs of each program.

DM often utilizes Regulatory Compliance Scales to determine the severity of non-compliance and guide the level of monitoring needed.

Integrative Monitoring Systems (IMS):

These systems go beyond simply checking compliance and aim to assess the overall quality of a program.

They integrate data from various sources, including IBPM, DM, and other program-specific metrics, to provide a holistic picture of performance.

IMS can inform licensing decisions by considering not only compliance but also program effectiveness in achieving its goals.

Here's a simplified analogy to illustrate the relationships:

Think of regulations as traffic rules.

IBPM is like a police officer checking every car for every violation, regardless of severity.

DM is like a police officer focusing on patrolling areas with high accident rates or known reckless drivers.

Regulatory Compliance Scales are like different levels of fines based on the severity of the traffic violation.

IMS is like a traffic management system that collects data on accidents, traffic flow, and road conditions to optimize traffic flow and safety.

Relationships:

RCS forms the foundation for DM and IMS by providing a way to assess degrees of compliance.

IBPM provides data for RCS and can be incorporated (with adaptations) into DM and IMS.

DM builds on RCS and IBPM by differentiating the intensity of monitoring based on risk and compliance.

IMS is the most comprehensive approach, integrating RCS, IBPM, DM, and additional data sources for a deeper understanding of program performance.

Regulatory Compliance Scales can be used within any of the monitoring approaches to provide a more nuanced assessment of compliance.

IBPM can be a starting point for differential monitoring, providing data on rule compliance to inform risk assessments.

Differential monitoring can be integrated into an integrative monitoring system, along with other data sources, to provide a comprehensive picture of program performance.

Here are some additional points to consider:

The choice of the most appropriate approach will depend on the specific context, such as the type of program being regulated and the available resources.

Implementation of these alternative paradigms requires careful planning and training of regulators and program providers.

Ongoing research and evaluation are needed to refine these approaches and ensure their effectiveness.

These alternative paradigms offer a more flexible and effective approach to licensing decisionmaking compared to the traditional IBPM approach. They allow for a better understanding of program strengths and weaknesses, optimize resource allocation, and ultimately lead to better regulatory outcomes.

These concepts offer a shift from traditional "one-size-fits-all" compliance models to more flexible and nuanced approaches that consider risk, program quality, and degrees of compliance. This can lead to more efficient and effective regulatory systems that support program improvement while protecting public safety.

Ultimately, these concepts offer alternative paradigms for licensing decision-making, moving away from a rigid "one-size-fits-all" approach to a more nuanced and risk-based system that considers both compliance and program quality.

References:

Fiene, R. (2019). A treatise on Regulatory Compliance. ***Journal of Regulatory Science, Volume 7***, 2019. https://doi.org/10.21423/jrs-v07fiene

Fiene (2022a). Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary. (2022). ***Journal of Regulatory Science, 10*(1). https://doi.org/10.21423/JRS-V10A239**

Fiene (2022b). Regulatory Compliance Scale, ***RIKINotes Blog***, January 2022.

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

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

Fiene (2025a). Finding the Right Rules. ***American Scientist, Volume 113, 1***. pps 16-19.

Fiene (2025b). The Uncertainty-Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies, ***Knowledge Journal***, under review.

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

Richard Fiene PhD, Senior Research Psychologist, Research Institute for Key Indicators (RIKILLC); Professor of Psychology (ret), Penn State University; Senior Research Consultant, National Association for Regulatory Administration (NARA). Email: Rfiene@RIKInstitute.com Website: http://RIKInstitute.com