

Potential Solution to the Child Care Trilemma Revisited

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

by Richard Fiene

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

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

of North Carolina at Greensboro and there were discussions about the revision to the Federal Interagency Day Care Requirements (FIDCR). What was the right mix, the balance of protections and quality enhancements for young children in child care that the Federal Department of Health, Education and Welfare wanted to promulgate nationally.

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

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



Rick Fiene
Exchange Leadership Initiative
exchangeexpress.com/leadership

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

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

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

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

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

to make monitoring more effective and efficient by focusing on risk to children and prediction of overall regulatory compliance (Fiene, 2019, 2025).

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

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

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

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

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

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

The point of this research abstract position paper is for us to take a step back and avoid a knee-jerk reaction to dealing with the child care crisis and that the only solution is to increase availability and affordability at the expense of health, safety and quality via deregulation (NAEYC, 2024)(Hechinger Report, 2024). We now have an emerging regulatory science

(Fiene, 2025) to guide us and I hope we use it for making educated and informed choices as we move forward in attempting to solve the continuing child care trilemma.

References

- Abt Associates. (1978). *Day Care Centers in the U.S.A National Profile 1976-1977*, Boston, Massachusetts.
- American Academy of Pediatrics and the American Public Health Association and the National Resource Center for Health and Safety in Child Care: AAP, APHA, NRCHSCC. (2019). *Caring for Our Children: The National Health and Safety Performance Standards*. Itasca, Illinois.
- Administration for Children and Families: ACF. (2015). *Caring for our Children Basics*. Washington, DC.
- Fiene, R. (2025). Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *American Scientist*, 113 (January-February 2025).
- Fiene, R. (2024). *Regulatory Compliance and Monitoring Systems*. Research Institute for Key Indicators and National Association for Regulatory Administration.
- Fiene, R. (2019). A Treatise on the Theory of Regulatory Compliance. *Journal of Regulatory Science*, 7, 1-3.
- Fiene, R. (1997). Searching for a Solution to the Child Care Trilemma. *Child Care Information Exchange*, September, 1997, 57-59, Redmond, Washington.
- Fiene, R. & Stevens, S. (2021). *Contact Hours as a New Metric Replacing Group Size and Staff-Child Ratios as well as a New Metric for COVID19 Thresholds*. Research Institute for Key Indicators/Penn State.
- Hechinger Report. (2024). *Investigating How Kids are Harmed by Lower Child Care Standards*. New York, New York.
- National Resource Center for Health and Safety in Child Care: NRCHSCC. (2019). *Stepping Stones to Caring for our Children*. Boulder, Colorado.
- National Association for the Education of Young Children: NAEYC. (2024). *The Costs of Deregulating Child Care*. Washington, DC.



knowledge

Study Protocol

The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies

Richard Fiene



Study Protocol

The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies

Richard Fiene 

Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University,
University Park, PA 16802, USA; rfiene@rikinstitute.com

Abstract: This research article proposes the use of an uncertainty–certainty matrix (UCM) for licensing decision making in the human services, which is the decision to issue a license to operate. It is a proposed study protocol and conceptual framework; it is not an empirical study. It shows how the matrix can be used in rule decision making and how it clearly shows when decision making has gone awry when bias is introduced into the decision making. It is also proposed to be used to make decisions in differential monitoring and in validation and reliability studies. This proposal presents a potential blueprint on how the UCM can be used within human services licensing as a decision-making tool.

Keywords: decision making; uncertainty–certainty matrix; regulatory compliance; licensing; reliability and validation studies

1. Introduction

This research proposal takes the Contingency Table, which is a well-known metric in the statistical decision-making research literature [1], and refocuses it on regulatory science within the context of the definition of regulatory compliance and licensing measurement. It also deals with the policy implications of this particular metric. In this study protocol, it is proposed that the Uncertainty–Certainty Matrix (UCM) is a fundamental building block to licensing decision making from a measurement perspective. The Contingency Table, as demonstrated by a 2×2 matrix, is utilized in regulatory compliance and is the center piece for determining licensing key indicator rules [2], but it is also a core conceptual framework in licensing measurement and ultimately in program monitoring and reviews [3].

The reason for selecting this matrix is the nature of licensing data: it is binary or nominal in measurement. Either a rule/regulation is in compliance or out of compliance. Presently, most jurisdictions deal with regulatory compliance measurement in this nominal level or binary level. There is to be no gray area; this is a clear distinction in making a licensing decision about regulatory compliance. The UCM also takes the concept of Inter-Rater Reliability (IRR) a step further in introducing an uncertainty dimension that is very important in licensing decision making which is not as critical when calculating IRR. Inter-Rater Reliability is a real concern in the human services licensing field in that in many cases it is difficult for jurisdictions to maintain a high degree of consistency when comparing individual licensing inspectors to each other. Part of the problem is a fundamental measurement issue; it is hoped that the addition of the UCM will help to mitigate this problem [4]. Licensing measurement is dominated by nominal measurement: either a rule is in compliance or it is out of compliance. A proposal has been suggested in which an ordinal scale based upon licensing rule violations would be utilized called the



Academic Editor: Cengiz Kahraman

Received: 1 January 2025

Revised: 25 April 2025

Accepted: 27 April 2025

Published: 28 April 2025

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

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Regulatory Compliance Scale (RCS) [3]. This new RCS scale shows promise, but it needs additional validation studies in order to be used on a regular basis for making human services licensing decisions (2a).

The reason for even suggesting this matrix is the high level of dissatisfaction with the levels of reliability in the results of program monitoring reviews as suggested in the previous paragraph. If the dissatisfaction was not so pronounced, it would not be an issue, but with it being so high, the field of licensing needs to take a proactive role in determining the best possible way to deal with increasing inter-rater reliability among licensing inspectors. Hopefully, this organizational schema via the UCM Matrix will help to think through this process related to licensing measurement and monitoring systems. The author has been conducting regulatory compliance studies for the past 50 years and has determined that the validity and reliability of these studies needs a more robust model for making licensing decisions via more accurate measurements of regulatory compliance. This led to the creation and proposing of the UCM Metric [5–7].

Over the past 50 years, it has been well documented by the National Association for Regulatory Administration (NARA) how the licensing field has changed in moving from a one-size-fits-all licensing and monitoring approach to one of differential or targeted licensing and monitoring (<https://www.naralicensing.org/key-indicators>, accessed on 24 April 2025). NARA has led this transition in the human services licensing and regulatory administration field, which has produced a much more productive, effective, and efficient licensing inspection system. The UCM and RCS are the latest pieces in the puzzle to accomplishing this new licensing decision-making framework.

The key pieces to the UCM are the following: the decision (D) regarding regulatory compliance and actual state (S) of regulatory compliance. Regulatory Compliance of individual Rules: Plus (+) = In-compliance, or Minus (−) = Out of compliance. As such, the matrix can be built as follows (Table 1):

Table 1. Uncertainty–Certainty Matrix (UCM) Logic Model.

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

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

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

2. Uncertainty–Certainty Matrix for Validation and Reliability Studies

This part of the research proposal is to explore the possibility of utilizing the Uncertainty–Certainty Matrix (UCM) as depicted in Table 1 in validation and reliability studies in licensing decision making. The UCM has been proposed for use in licensing decision making, but this would be an extension of this thinking to studies that involve validating licensing decisions, such as when key indicators/predictor rules are used in comparison with comprehensive reviews of rules [5] and in reliability studies to determine individual inspector bias in regulatory compliance [8,9].

The basic premise of the UCM is that individual decision making matches reality. When it comes to regulatory compliance decision making, a 2×2 matrix can be drawn with the possible outcomes as indicated in the following table (Table 2), which is based upon the logic of Table 1.

Table 2. Uncertainty–Certainty Matrix (UCM) Logic Model applied to Validation Studies.

UCM Matrix Logic	For Validation Studies	Decision Regarding	Regulatory Compliance
		(+) In Compliance	(−) Not In Compliance
Actual State of Compliance	(+) In Compliance	Agreement (++)	Disagreement (+−)
	(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

In using this table, the hope is that the decision regarding regulatory compliance matches the actual state of compliance where the coefficient is as close to +1.00 as possible; in other words, perfect agreement. The agreement cells are heavily weighted (++) and (−−). We do not want to see all the cells, both agreement and disagreement cells, equally weighted (++) , (+−), (−+), (−−). That would indicate a random response rate and a coefficient close to 0.00.

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

The UCM can be used for both reliability and validity testing as suggested in the above table (Table 2). For validity, false positives (+−) and negatives (−+) should either be eliminated or reduced as well as possible, and the remaining results should show the typical diagonal pattern as indicated by the agreement cells.

For reliability, the same pattern should be observed as in the validity testing above, but there is an additional test in which bias is tested for. Bias is ascertained if the patterns in the results indicate a horizontal or vertical pattern in the data with little or no diagonal indication. Bias can be found at the individual inspector level, as well as at the standard level or the actual state of compliance. This could provide a helpful visual for licensing administrators regarding how decisions are being made about program regulatory compliance in the field.

In both reliability and validity testing, random results in which each of the cells are equally filled are not desirable either. Obviously, additional training involving licensing inspectors would need to occur in order to make the data collection efforts both reliable and valid. Monitoring of regulatory compliance history data would need to occur on an ongoing basis to make sure that biases did not return or if new biases developed within the regulatory compliance system.

The following Tables 3–8 depict the above relationships with results highlighted in red:

Table 3. Valid and Reliable Results.

Valid and Reliable Results	(+) In Compliance	(−) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+−)
(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

Table 4. Random Results.

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

Table 5. Positive Bias Results Individual Assessor.

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

Table 6. Negative Bias Results Individual Assessor.

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

Table 7. Positive Bias Results Standard.

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

Table 8. Negative Bias Results Standard.

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

Tables 3–8 demonstrate the different results based upon individual response rates when making regulatory compliance decisions about rules. Table 3 is what needs to be attained and Tables 4–8 need to be avoided. Only in Table 3 are false negatives and positives eliminated or avoided. In Tables 4–8, false negatives and/or false positives are introduced, which is not desirable when making validity or reliability decisions.

Table 4 results clearly indicate that a great deal of randomness has been introduced in the regulatory compliance decision making in which the individual licensing inspector decisions do not match reality. Tables 5 and 6 demonstrate bias in the decision-making process either positively (inspector always indicates in compliance) or negatively (inspector always indicates out of compliance). It is also possible that the standard being used has bias built into it; this is less likely but is still a possibility. The results in Tables 7 and 8 demonstrate where this could happen.

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

3. Uncertainty–Certainty Matrix for Differential Monitoring Studies

The purpose of this part of this research proposal is to explore the possibility of utilizing the Uncertainty–Certainty Matrix (UCM) not only in validation and reliability studies in licensing decision making, but also with differential monitoring studies. The UCM has been proposed for use in licensing decision making, but this would be an extension of this thinking to studies that involve validating licensing decisions, such as when key indicators are used in comparison with comprehensive reviews of rules and in the development of risk rules as part of the risk assessment methodology [4]. This new Differential Monitoring 2×2 Matrix can also be used to depict the relationship between full and substantial regulatory compliance and the nature of rulemaking.

The basic premise of the DMM: Differential Monitoring Matrix is similar to the original thinking with the UCM Matrix Logic as depicted in Table 1, but there are some changes in the formatting of the various cells in the matrix (see Table 9). When it comes to regulatory compliance decision making, a 2×2 matrix can be drawn with the possible outcomes as is indicated in Table 9 where each individual rule is either in (+) or out (−) of compliance. Additionally, there is the introduction of a high regulatory compliant group (+) and a low regulatory compliant group (−), which is different from the original UCM.

Table 9. DMM—Differential Monitoring Matrix.

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

By utilizing the format of Table 9, several key components of differential monitoring can be highlighted, such as key indicators and risk assessment rules, as well as the relationship between full and substantial regulatory compliance.

Regulatory compliance is grouped into a high group (+); generally, this means that there is either full or substantial regulatory compliance with all rules. The low group (−) usually has 10 or more regulatory compliance violations [4]. Individual rules being in (+) or out (−) of regulatory compliance is self-explanatory.

Tables 10–16 below demonstrate the following relationships:

Table 10. Key Indicators/Predictor Rules.

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

Table 11. Risk Rules/Place Clients at Increased Risk.

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

Table 10 depicts the key indicator relationship between individual rules and the high/low groups as indicated in red. In this table, the individual rule is in compliance with

the high group and is out of compliance with the low group. This result occurs on a very general basis and should have a 0.50 coefficient or higher with a p value of less than 0.0001.

Table 12. Full Compliance with All Rules.

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

Table 13. Substantial Compliance with All Rules.

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

Table 14. Very Difficult Rules.

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

Table 15. Poor Performing Programs.

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

Table 16. Terrible Rule.

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

Table 11 depicts what most rules look like in the 2×2 DMM. Most rules are always in full compliance since they are standards for basic health and safety for individuals. This is especially the case with rules that have been weighted as high-risk rules. Generally, one never sees non-compliance with these rules. There will be a substantial number of false positives (+−) found with high-risk rules, but that is a good thing.

Table 12 depicts what happens when full compliance is used as the only criterion for the high group. Notice that the cell right below (++) is eliminated (−+). This is highly recommended since it eliminates false negatives (−+) from occurring in the high group. As is seen in Table 12, when substantial compliance is used as part of the high group sorting, false negatives are re-introduced. If possible, this should be avoided; however, in some cases, because of the regulatory compliance data distribution, this is not always possible where not enough full compliant programs are present.

Table 13 depicts what occurs when substantial compliance is used as part of determining the high group. False negatives can be reintroduced into the matrix which needs to be either eliminated or reduced as best as possible. If substantial compliance needs to be used in determining the high group, then there is a mathematical adjustment that can be made, which will impact the equation and essentially eliminate false negatives mathematically.

Table 14 depicts what happens if the individual rule is particularly difficult to comply with. Both the high performers as well as the low performers are out of compliance with the rule.

Table 15 depicts a situation where the programs are predominantly in a low group with few at full or substantial regulatory compliance, which is indicative of poor performing programs. Very honestly, this is generally not seen in the research literature, but it is a possibility and one to be in tune with.

Table 16 depicts a terrible individual rule which predicts just the opposite of what we are trying to do with programs. Obviously, this rule would need to be rewritten so that it fits with the essence of regulatory compliance in helping to protect individuals.

The following Tables 10–16 depict the above relationships with results highlighted in red.

Tables 10–16 demonstrate the different results based on the relationship between individual regulatory compliance and if a program is either a high performer or a low performer. These tables are provided as guidance for understanding the essence of differential monitoring and regulatory compliance, which has various nuances when it comes to data distribution. This research proposal for a UCM hopefully can be used as a guide in determining from a data utilization point of view how to make important regulatory compliance policy decisions, such as which rules are excellent key indicator rules, which are performing as high risk rules, the importance of full compliance, what to do when substantial compliance needs to be employed, are there difficult rules to comply with, how well are programs performing, and do we have less than optimal rules that are in need of revision.

4. Conclusions

The Uncertainty–Certainty Matrix (UCM) should provide a useful tool for assessing the effectiveness of licensing decision making in the human services via validation and reliability studies within differential monitoring systems by visually inspecting cell proportions to determine if the appropriate results are depicted in the above matrices.

It is hoped that licensing researchers and regulatory scientists will experiment with it and test it out in different arenas beyond early care and education programs. It appears to have broad applicability across regulatory disciplines. The matrix has helped to identify the need to address false positives and negatives in the human services licensing decision-making process which undermines the effort of protecting clients.

The UCM also appears to provide a framework to identify reliability issues across licensing inspectors carrying out evaluations of individual programs. This issue of reliability is a big issue in the human services licensing field where there is a great deal of inconsistency when it comes to measuring regulatory compliance [10–12]. The UCM could be applied to existing regulatory compliance history data to determine if bias is present or not. It provides a clear visual demonstration of when regulatory compliance history data have gone awry and are not performing as they should. This can be a useful tool for licensing administrators in making changes to their overall licensing system, as well as for which individual rules/regulations/standards are most effective in protecting clients or might need revision.

The major limitation of the UCM is that as of this writing, it has not been empirically tested to see if this conceptual framework is really helpful to licensing policymakers and researchers. The UCM is a theoretical model at this point that needs to be verified. At the same time, it holds promise for the human services licensing field because the field as it relates to regulatory science has a measurement problem when it comes to reliability and

validity. Without a solid measurement structure, it is the old adage of “Garbage In, Garbage Out”. Hopefully, the UCM is a first step to rectifying this issue.

Clearly, for future research, there needs to be additional expansion beyond the child-care and early education field to all of human services and then beyond this scope to other regulatory areas to determine if a UCM approach is relevant. It is obvious that in clinical studies within the medical field that the UCM would be very appropriate in order to avoid false negatives where a drug’s side effects would be more detrimental than the potential benefits from taking the particular drug. We need additional real-life examples where the UCM model can be tested to see how useful it would be in other regulatory settings.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pearson, K. On the Theory of Contingency and Its Relation to Association and Normal Correlation. In *Draper’s Company Research Memoirs, Biometric Series I*; Cambridge University Press: London, UK, 1904.
2. Fiene, R. A Treatise on Regulatory Compliance. *J. Regul. Sci.* **2019**, *7*, 1–3. [[CrossRef](#)]
3. Fiene, R. *The Development of a Regulatory Compliance Scale*; Research Institute for Key Indicators/Penn State University: University Park, PA, USA, 2025.
4. Miranda, B.; Ekyalongo, Y.; Franchett, A.; Maxwell, K. Monitoring Practices Used in Child Care and Early Education Licensing. In *OPRE Report #2022-137*; Office of Planning, Research, and Evaluation, Administration for Children and Families, Department of Health and Human Service: Washington, DC, USA, 2022.
5. Fiene, R. *Regulatory Compliance*; National Association for Regulatory Administration: Fredericksburg, VA, USA, 2024.
6. Trivedi, P.A. *Innovation in Monitoring in Early Care and Education: Options for States*; Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services: Washington, DC, USA, 2015.
7. Fiene, R. Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *Am. Sci.* **2025**, *113*, 16–21.
8. Maxwell, K.; Miranda, B.; Harris, P.; Johnson, N.; Padilla, C.; Banghart, P.; Verhoye, A.; Early, D.M. Child care and early education licensing literature review. In *OPRE Report #2024-199*; Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA, 2024.
9. Zellman, G.; Fiene, R. Validation of quality rating and improvement systems for early care and education and school-age care. In *OPRE and Child Trends*; Office of Planning Research and Evaluation Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA, 2012.
10. Lahti, M.; Elicker, J.; Zellman, G.; Fiene, R. Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs. *Early Child. Res. Q.* **2014**, *30*, 280–290. [[CrossRef](#)]
11. Fiene, R.; Kroh, K.E. *Human Services Licensing Measurement, Regulatory Compliance, and Program Monitoring Systems*; Research Institute for Key Indicators & the National Association for Regulatory Administration: Minneapolis, MN, USA, 2016.
12. Payne, A.L. *Strong Licensing: The Foundation for a Quality Early Care and Education System*; National Association for Regulatory Administration: Minneapolis, MN, USA, 2011.

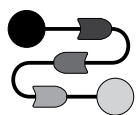
Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

A reprint from

American Scientist

the magazine of Sigma Xi, The Scientific Research Honor Society

This reprint is provided for personal and noncommercial use. For any other use, please send a request to Permissions, American Scientist, P.O. Box 13975, Research Triangle Park, NC, 27709, U.S.A., or by electronic mail to perms@amsci.org.
©Sigma Xi, The Scientific Research Honor Society and other rightsholders



Finding the Rules that Work

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

Richard Fiene

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

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

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

Needless to say, developmentally appropriate curricula change kids' lives; boasting a perfect record does not. This observation neither dismisses

the 200 to 400 rules and regulations set by respective U.S. states nor undermines the importance of complying with them, either as individual rules or in the aggregate. And full compliance does improve safety. But, as data gathered by my research team repeatedly demonstrates, a vague, uncomfortable gap separates full, costly regulatory compliance from program quality.

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

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

Such a discussion is long overdue. The unequal worth of many general licensing and quality standards, including those driven by a regulatory political

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

Modernizing Measurement

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

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

QUICK TAKE

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

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

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



DGLimages/Shutterstock

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

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

The watershed moment for regulatory science as it pertains to children's programs came in the 1980s. The previous decade's major childcare expansion in the United States had created a backlog of licensing assessments, caused unmanageable monitoring delays, and laid bare the logistical limits of case studies. These factors, combined with advances in computing, led states to introduce an empirical, quantitative, and instrument-based approach, complete with sophisticated software systems designed by state

agencies and private vendors to track regulatory compliance and quality assessment data. Empirical evidence not only moved regulatory science from qualitative to quantitative analysis, it also revealed surprising patterns.

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

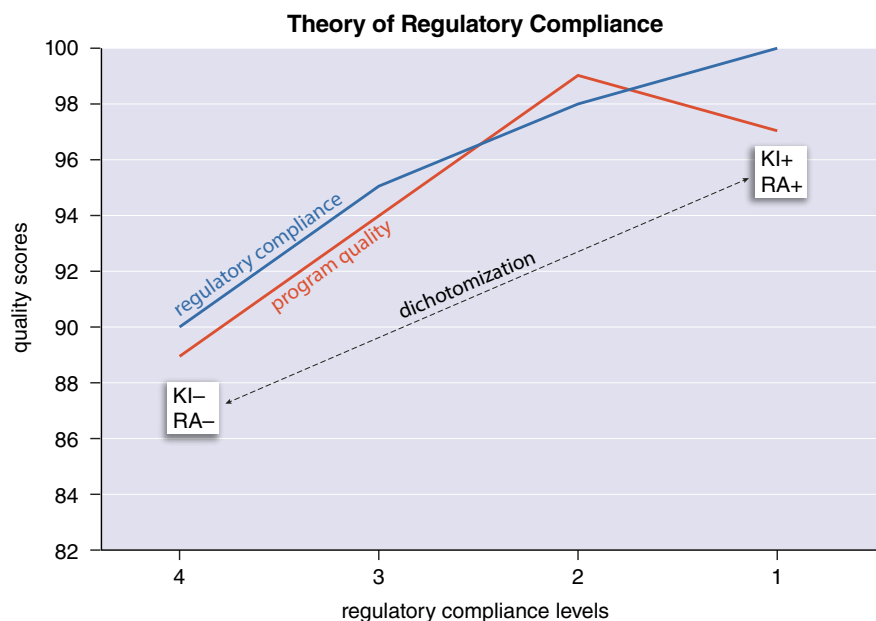
But as expert opinion and anecdotal evidence gave way to better-designed studies and empirical data, and as

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

A New Paradigm

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

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



adapted from Richard Fiene

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

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

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

tions, training, or technical assistance, or may undergo license suspension or even permanent closure.

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

Here, and in my research, I primarily deal with center-based care programs, but the findings apply to other service types as well, such as family childcare homes and school-age programs, as well as human services categories such as child residential, child foster care, adult residential, and adult personal care homes. My data and research concern the relationship between quality and compliance, and how to improve it. They stem from studies of hundreds of programs I conducted at the state level

from the 1970s through the 2010s, when I directed various research and training institutes at Pennsylvania State University. In these controlled and replicated studies, trained observers collected both regulatory data and program quality data from eight states, three Canadian provinces, and the U.S. Head Start program. The work ran the gamut, from site selection via stratified random samples, to dispatching data collectors to specific programs, to providing individual states with an overall blueprint describing how to conduct their studies.

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

Differential Monitoring

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

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

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

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

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

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

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

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

adapted from Richard Fiene

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

targeted, and focused reviews take approximately 50 percent less time than comprehensive reviews.

Unfortunately, although many licensing bodies use risk assessment or key indicator methodologies, few use both. *Monitoring Practices Used in*

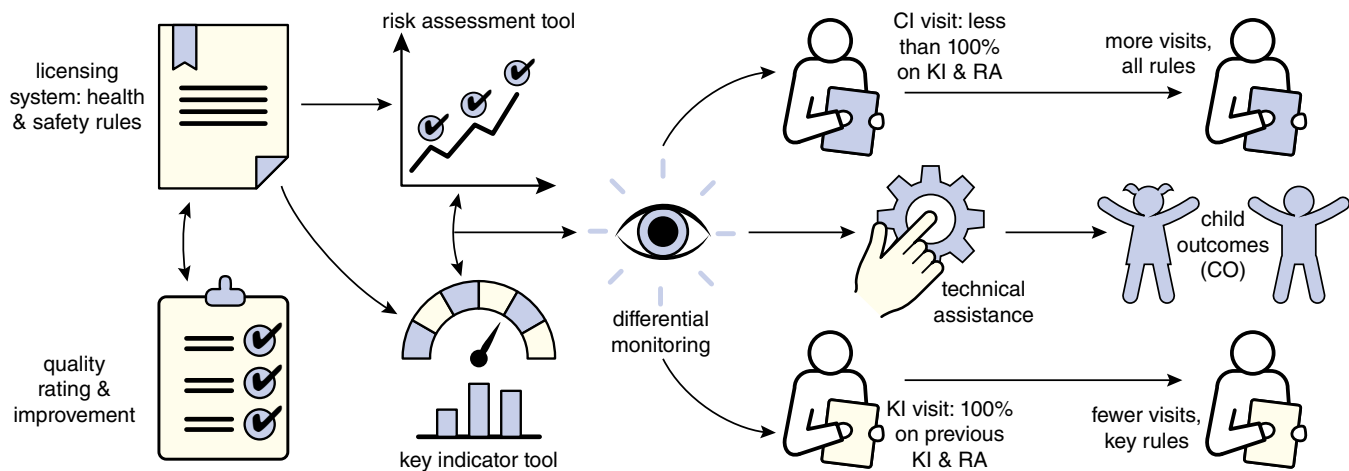
If, as data suggested, substantially compliant programs provided the same or better care as fully compliant ones, then clearly we needed to rethink our program evaluation strategies.

Child Care and Early Education Licensing, a federal accounting of how states conduct program monitoring, reported that 10 states used key indicators, 17 states used risk assessments, and only one state used both. Hopefully, this pattern will change as the regulatory science field matures over the coming decades.

Since I first proposed it in the mid-1980s, the theory of regulatory compliance has faced numerous critics in the human services licensing field, especially among advocates of uni-

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

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



Barbara Aulicino

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

Rethinking Nominal Data

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

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

Let’s begin with the latter question. In an ideal world, judgments made by assessors would perfectly reflect a program’s actual regulatory compliance state. But research that tests reliability and replicability in the licensing

field empirically shows a concerning degree of disagreement when a second observer validates the decision regarding regulatory compliance. These disagreements suggest a worrying number of false positives and false negatives.

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

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

multicategory data, we must first dichotomize it into two distinct groups.

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

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

Infusing Quality

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

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

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

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

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

Should those findings bear out the value of evaluating individual rules via the 1–7 regulatory compliance scale, I propose that it should contain the fol-

lowing categories: exceeding full compliance, full compliance, substantial compliance, and mediocre compliance (see figure on page 19). These categories differ from the aggregate rule compliance scale currently under evaluation (full, substantial, mediocre, and low compliance) because aggregate compliance only considers health and safety elements, whereas an individual scale would also take quality into account.

Research supports the value of transitioning from uniform monitoring and full compliance to differen-

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

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

Looking Forward

The regulatory compliance scale is a new and evolving metric. It transforms licensing data from a mere violation tally into a more useful and intuitive scale, one more consistent with the program quality measurements supported by research. Hereafter, I hope that the approach will incorporate quality measurements and more nuanced weighting into the evaluation of individual rule compliance. But dis-

cussions are just beginning, and this shift will pose a substantial challenge for agencies, which must also cope with the aftermath of the COVID-19 pandemic and a rising tendency toward deregulation.

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

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

Bibliography

- Fiene, R. 2023. Licensing measurement and monitoring. National Association for Regulatory Administration licensing curriculum. Published online www.naralicensing.org/key-indicators
- Fiene, R. 2022. Regulatory compliance monitoring paradigms and the relationship of regulatory compliance/licensing with program quality: A policy commentary. *Journal of Regulatory Science* 10:1–7.
- Fiene, R. 2019. A treatise on regulatory compliance. *Journal of Regulatory Science* 7:1–3.

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

Structural and Process Quality in Early Childhood Education

Subjects: **Psychology**

Contributor: Richard Fiene

This article demonstrates the contribution that the theory of regulatory compliance makes as a unifying framework for child care and early education structural and process quality.

early childhood education, regulatory compliance, regulatory science, structural quality, process quality, child care

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

1. Introduction

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.

2. 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

3. 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.

4. 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.

Retrieved from <https://encyclopedia.pub/entry/history/show/130384>

Licensing Measurement and Monitoring Systems

Subjects: **Psychology**

Contributor: Richard Fiene

In the realm of human services regulatory administration, ensuring compliance with licensing requirements is crucial for maintaining quality standards and safeguarding the well-being of individuals receiving care. As regulatory agencies strive to enhance their oversight and monitoring capabilities, the integration of measurement and monitoring systems has emerged as a valuable tool. This paper explores the significance of licensing measurement and monitoring systems and delves into the application of regulatory science in the context of human services regulatory administration. It will deal with several issues related to this topic and expand its content beyond early care and education which has been more of the focus previously.

licensing measurment

monitoring systems

human services

Licensing Measurement and Monitoring Systems: Regulatory Science Applied to Human Services Regulatory Administration

Richard Fiene PhD¹

June 2023

In the realm of human services regulatory administration, ensuring compliance with licensing requirements is crucial for maintaining quality standards and safeguarding the well-being of individuals receiving care. As regulatory agencies strive to enhance their oversight and monitoring capabilities, the integration of measurement and monitoring systems has emerged as a valuable tool.

This paper explores the significance of licensing measurement and monitoring systems and delves into the application of regulatory science in the context of human services regulatory administration. It will deal with several issues related to this topic and expand its content beyond early care and education which has been more of the focus previously.

Licensing measurement and monitoring systems play a crucial role in regulatory administration for several reasons:

- **Compliance Verification:** Regulatory agencies need to ensure that businesses and individuals comply with specific laws, regulations, and standards. Licensing measurement and monitoring systems provide a means to verify compliance by collecting data and measuring various parameters. These systems help regulators

determine whether license holders are meeting the required standards and taking appropriate actions to mitigate risks.

- **Quality Assurance:** Licensing measurement and monitoring systems contribute to quality assurance efforts by assessing the performance of licensed entities. They enable regulators to monitor the quality of services and activities associated with the licensing process. By establishing measurement criteria and tracking the relevant metrics, regulators can ensure that license holders maintain the desired level of quality and meet the expectations of consumers or the public.
- **Risk Management:** Many industries involve inherent risks that need to be managed effectively. Licensing measurement and monitoring systems allow regulatory agencies to assess and monitor the risks associated with licensed activities. By continuously monitoring key indicators, regulators can identify potential risks, deviations from safety standards, or non-compliance issues. This information helps regulators take appropriate actions to minimize risks and ensure public safety.
- **Data-Driven Decision Making:** Licensing measurement and monitoring systems generate substantial amounts of data that can be analyzed to make informed decisions. Regulators can analyze trends, patterns, and performance metrics to identify areas of concern or improvement. Data-driven insights enable regulators to make evidence-based decisions, allocate resources effectively, and prioritize enforcement actions where they are most needed.
- **Enforcement and Remediation:** When non-compliance or deviations from regulations are identified, licensing measurement and monitoring systems provide evidence to support enforcement actions. Regulators can use the data collected to take appropriate enforcement measures, such as issuing warnings, imposing penalties, or revoking licenses. These systems also help in tracking the progress of remedial actions taken by license holders to address any identified issues or deficiencies.
- **Transparency and Accountability:** Licensing measurement and monitoring systems enhance transparency and accountability in regulatory administration. By implementing these systems, regulators can demonstrate their commitment to fair and consistent enforcement of regulations. The data collected and analyzed can be made accessible to the public, stakeholders, and policymakers, fostering trust, and allowing for external scrutiny of regulatory processes.

Licensing measurement and monitoring systems are vital in regulatory administration as they facilitate compliance verification, quality assurance, risk management, data-driven decision making, enforcement, and accountability. These systems help regulators ensure that licensed entities operate within the set standards, mitigate risks effectively, and safeguard the interests of the public.

Regulatory Science is relevant to human services regulatory administration in all industries. Regulatory science is the scientific discipline that combines various fields, including law, public policy, data analysis, and risk assessment, to inform and guide regulatory decision-making.

Measurement and monitoring systems are regulatory science aids in the development and implementation of evidence-based regulations and policies.

Regulatory agencies overseeing a wide range of human services, such as healthcare facilities, child care centers, mental health institutions, and more, face several challenges in their oversight role. Some of the key challenges include:

- **Diverse and Complex Landscape:** The human services sector encompasses a broad range of industries, each with its unique complexities, regulations, and standards. Regulatory agencies must navigate and understand this diverse landscape to effectively oversee and enforce compliance. The sheer variety of services, settings, and stakeholders involved makes it challenging to develop uniform regulations and monitoring approaches that address the specific needs of each sector.
- **Rapidly Evolving Practices and Technologies:** The human services field is constantly evolving, with new practices, technologies, and treatments emerging. Regulatory agencies need to keep pace with these changes to ensure that the regulations remain relevant and up-to-date. However, this can be a challenging task, as it requires continuous monitoring, research, and adaptation of regulations to address emerging risks and advancements adequately.
- **Resource Constraints:** Regulatory agencies often face resource constraints in terms of staffing, funding, and technological capabilities. Insufficient resources can limit their capacity to conduct thorough inspections, investigations, and monitoring activities. Additionally, limited resources may also impact the frequency and intensity of oversight, making it difficult to identify and address compliance issues effectively.
- **Compliance Variability:** Human services facilities and institutions can vary significantly in terms of size, ownership, resources, and compliance history. Regulatory agencies need to develop oversight strategies that account for these variations while ensuring consistent enforcement and quality standards across the board. Balancing the need for flexibility with the need for uniformity is a constant challenge for regulatory agencies. And this becomes increasingly complex when dealing with the regulatory compliance theory of diminishing returns/ceiling effect.
- **Stakeholder Engagement and Resistance:** Regulatory oversight often involves engaging with various stakeholders, including facility owners, professionals, service recipients, advocacy groups, and the public. These stakeholders may have different interests, priorities, and perspectives, leading to potential conflicts or resistance to regulatory measures. Balancing the diverse viewpoints and managing stakeholder expectations is essential for effective oversight.
- **Data Management and Analysis:** The vast amount of data generated by human services facilities can pose challenges in terms of data management, analysis, and interpretation. Regulatory agencies need robust systems and processes to collect, store, analyze, and make sense of the data to identify trends, patterns, and areas of concern. The integration and interoperability of data systems across different sectors and agencies can be complex and time-consuming.
- **Legal and Ethical Considerations:** Regulatory agencies must operate within legal frameworks and adhere to ethical standards while overseeing human services. They need to strike a balance between protecting public health and safety and respecting individual rights and privacy. Navigating legal complexities, ensuring due process, and maintaining confidentiality can be challenging in an environment where ethical dilemmas may arise.

Addressing these challenges requires a proactive and adaptive approach from regulatory agencies. They need to foster collaboration with stakeholders, invest in capacity-building efforts, leverage technology for efficient data management, and engage in continuous evaluation and improvement of their oversight strategies.

Inadequate monitoring in the human services can have significant risks and consequences, highlighting the need for robust systems that ensure compliance and promote accountability. Human services encompass a wide range of sectors, including healthcare, social welfare, child protection, and criminal justice. Monitoring in these areas is essential to safeguard the well-being and rights of individuals, prevent abuses, and ensure the effective delivery of services. Here are some potential risks and consequences of inadequate monitoring:

- **Abuse and neglect:** Without proper monitoring, vulnerable individuals may be at a higher risk of abuse, neglect, or exploitation. For instance, in healthcare settings, inadequate monitoring can lead to medical errors, mistreatment of patients, or substandard care. Similarly, in child protection services, insufficient monitoring can result in children remaining in abusive or neglectful environments.
- **Violation of rights:** Inadequate monitoring can lead to violations of individuals' rights, including their civil liberties, privacy, and dignity. For example, in criminal justice systems, inadequate monitoring can result in wrongful convictions, excessive use of force, or violations of prisoners' rights. In social welfare programs, lack of monitoring can lead to discrimination, improper denial of benefits, or infringement of recipients' rights.
- **Inefficiency and ineffective service delivery:** Monitoring is crucial for evaluating the effectiveness and efficiency of human services. Without robust monitoring systems, it becomes challenging to identify gaps, assess performance, and make informed decisions for improvement. Inadequate monitoring may lead to wastage of resources, duplication of efforts, or the continuation of ineffective programs that fail to meet the needs of the intended beneficiaries. This is where risk assessment rules and key indicator rules play an important role in increasing the effectiveness and efficiency of the monitoring process by utilizing a more differential monitoring approach.
- **Lack of accountability:** Monitoring plays a vital role in ensuring accountability within human service systems. It helps identify and address instances of misconduct, malpractice, or non-compliance with regulations and standards. Inadequate monitoring can result in a lack of transparency and accountability, allowing misconduct to go unnoticed, perpetrators to go unpunished, and systemic problems to persist.
- **Loss of public trust:** Inadequate monitoring erodes public trust in human service systems. When people perceive that their well-being, rights, or safety are compromised due to poor monitoring, it undermines their confidence in these services. Public trust is crucial for the effective functioning of human services, as it promotes cooperation, engagement, and participation of individuals and communities.

To mitigate these risks and consequences, robust monitoring systems are essential. Such systems should include clear guidelines, regular inspections, audits, reporting mechanisms, and independent oversight bodies. They should also leverage technology and data analysis to enhance monitoring capabilities and identify patterns or anomalies. Additionally, staff training on monitoring protocols and the establishment of a culture of accountability are crucial components of an effective monitoring framework.

Inadequate monitoring in human services poses significant risks and consequences. It can lead to abuse, neglect, rights violations, inefficiencies, lack of accountability, and loss of public trust. Robust monitoring systems, incorporating clear guidelines, regular inspections, technology, and independent oversight, are necessary to ensure compliance, protect individuals, and promote accountability within human service sectors.

The integration of measurement and monitoring systems into the licensing process in human services is a crucial development that leverages technology and data analytics to track, evaluate, and verify compliance with licensing standards. These systems provide real-time monitoring capabilities, enabling early detection of non-compliance, improved transparency, and enhanced accountability. Let's delve into the details of how these systems work and the benefits they bring.

Measurement and monitoring systems in the context of human services licensing involve the use of advanced technologies, such as sensors, cameras, electronic record-keeping systems, and data analytics tools. These technologies are integrated into the licensing process to collect, analyze, and interpret relevant data in real-time. The aim is to ensure that organizations and individuals providing human services comply with the established licensing standards and regulations.

One significant advantage of integrating measurement and monitoring systems is the early detection of non-compliance. With real-time monitoring, regulatory agencies can identify potential violations promptly. For example, if a human services facility is required to maintain a specific temperature range, sensors can continuously monitor the temperature levels. If there is a deviation from the acceptable range, an alert can be triggered, enabling swift corrective action. This early detection mechanism helps prevent potential risks and harm to individuals receiving those services.

Moreover, these systems improve transparency by providing accurate and objective data. Instead of relying solely on periodic inspections or self-reported information, regulatory agencies can access real-time data collected by the monitoring systems. This data-driven approach ensures a more comprehensive and accurate assessment of compliance with licensing standards. It reduces the reliance on subjective observations and minimizes the possibility of information gaps or bias.

Furthermore, integrating measurement and monitoring systems enhances accountability for organizations and individuals providing human services. By continuously monitoring and recording data, these systems create an audit trail that can be used for accountability purposes. The collected data provides evidence of compliance or non-compliance with licensing standards, which can be used in regulatory investigations or legal proceedings if necessary. This level of accountability fosters a culture of responsibility and incentivizes compliance with licensing requirements.

The benefits of these systems extend beyond regulatory agencies. Service providers themselves can benefit from real-time monitoring by gaining insights into their own operations and performance. By analyzing the data collected, they can identify areas for improvement, optimize resource allocation, and make evidence-based

decisions to enhance the quality of their services. This data-driven approach supports continuous improvement and helps providers meet and exceed licensing standards.

The integration of measurement and monitoring systems into the licensing process in human services offers significant advantages. It leverages technology and data analytics to enable real-time monitoring, early detection of non-compliance, improved transparency, and enhanced accountability. These systems provide regulatory agencies with objective data to ensure

compliance with licensing standards and promote the safety and well-being of individuals receiving human services. Simultaneously, service providers benefit from insights gained through data analysis, allowing them to optimize their operations and deliver higher quality services.

Licensing measurement and monitoring systems in human services play a crucial role in ensuring compliance with regulations, tracking licensing activities, and monitoring the quality and safety of services provided. These systems typically consist of several key components that work together to enable effective measurement and monitoring. Here are the main components:

- **Comprehensive Databases:** A central database is essential for storing all licensing-related information, including provider details, facility data, licensing standards, inspection reports, and compliance history. These databases provide a foundation for data collection, analysis, and reporting.

Example: The Child Care Licensing System (CCLS) developed by the Administration for Children and Families in the United States is a comprehensive database that tracks and manages child care licensing information. It allows agencies to manage licensing processes, track violations, and generate reports.

- **Automated Data Collection Tools:** Automation tools streamline the process of data collection by capturing information electronically, reducing manual effort, and improving accuracy. These tools can include online application forms, electronic submission of documentation, and automated notifications.

Example: The Integrated Regulatory Information System (IRIS) used by the California Department of Social Services enables online application submissions, digital document management, and automated notifications for licensing updates. It simplifies the data collection process and enhances efficiency.

- **Risk Assessment Algorithms:** Risk assessment algorithms help identify high-risk facilities or providers that require increased monitoring or intervention. These algorithms analyze various factors such as compliance history, complaint data, inspection results, and other relevant indicators to prioritize resources effectively.

Example: The Risk Assessment and Management Tool (RAM) implemented by the Australian Government's Department of Health is used to assess and manage risks associated with aged care services. RAM employs algorithms that analyze data on quality indicators, complaints, and non-compliance to determine risk levels and allocate resources accordingly.

- **Data Visualization Platforms:** Data visualization platforms present licensing data in a user-friendly and meaningful way, allowing regulatory agencies to monitor trends, identify patterns, and make data-driven decisions. These platforms often include interactive dashboards, charts, and reports.

Example: The Licensing Information System (LIS) developed by the Department of Health and Human Services in the state of Maine provides a data visualization platform that allows users to generate customized reports, view interactive charts, and track licensing compliance trends.

- **Compliance Monitoring Tools:** Compliance monitoring tools assist in conducting inspections, audits, and other monitoring activities efficiently. These tools can include mobile applications for inspectors to collect data on-site, electronic checklists, and automated scheduling of inspections.

Example: The Licensing Automation System (LAS) implemented by the Minnesota Department of Human Services offers mobile applications for licensing staff to perform inspections, record findings, and generate inspection reports on the go. It simplifies the monitoring process and improves accuracy.

Overall, these components work together to create effective licensing measurement and monitoring systems in human services. By leveraging comprehensive databases, automated data collection tools, risk assessment algorithms, data visualization platforms, and compliance monitoring tools, regulatory agencies can enhance their oversight capabilities, improve efficiency, and ensure the provision of high-quality services while maintaining compliance with regulations.

Licensing measurement and monitoring systems have had a significant impact on regulatory administration and the human services sector. These systems play a crucial role in enabling regulators to proactively identify potential risks, address compliance issues promptly, and ensure the safety and quality of services provided. In this response, we will discuss the impact of these systems and provide case studies and examples that illustrate the positive outcomes achieved through their implementation.

One of the primary benefits of licensing measurement and monitoring systems is their ability to provide regulators with real-time data and insights. These systems collect and analyze various metrics and indicators, allowing regulators to monitor the performance and compliance of service providers. By having access to accurate and up-to-date information, regulators can proactively identify potential risks and address them before they escalate into serious problems.

For instance, let's consider the case of a regulatory agency responsible for overseeing childcare facilities. By implementing a licensing measurement and monitoring system, the agency can track key indicators such as staff-to-child ratios, health and safety inspections, and educational programs. If the system detects any deviations from the established standards, it can alert regulators, enabling them to intervene promptly. This proactive approach helps prevent incidents and ensures that children receive appropriate care and support.

Another positive outcome of licensing measurement and monitoring systems is improved compliance management. These systems streamline the process of monitoring and assessing compliance with regulations and standards.

Service providers can input data directly into the system, reducing the administrative burden and ensuring accuracy. Regulators can then use this data to identify patterns, assess compliance levels, and take appropriate actions if non-compliance is detected.

For example, let's consider the case of a regulatory agency overseeing healthcare facilities. With a licensing measurement and monitoring system in place, the agency can track indicators such as medication errors, infection rates, and patient satisfaction scores. If the system identifies a healthcare facility with consistently high medication error rates, regulators can conduct targeted inspections and work closely with the facility to implement corrective measures. This proactive approach not only improves patient safety but also helps service providers enhance the quality of care they deliver.

Furthermore, licensing measurement and monitoring systems contribute to transparency and accountability in the human services sector. These systems provide a centralized platform where regulators, service providers, and the public can access information about licensing status, compliance records, and performance metrics. By promoting transparency, these systems help build trust among stakeholders and empower individuals to make informed decisions about service providers.

For instance, in the context of elder care services, a licensing measurement and monitoring system can provide a public database that includes information on the licensing status of assisted living facilities, compliance records related to safety standards, and ratings based on resident satisfaction surveys. This enables families and individuals seeking care for their loved ones to make informed choices and select facilities that meet their specific needs.

Licensing measurement and monitoring systems have had a transformative impact on regulatory administration and the human services sector. These systems enable regulators to proactively identify potential risks, address compliance issues promptly, and ensure the safety and quality of services provided. Through case studies and examples, we have seen how these systems have improved oversight in childcare, healthcare, and elder care, leading to positive outcomes such as enhanced safety, improved compliance, and increased transparency. The

implementation of such systems has the potential to further strengthen regulatory efforts and promote the well-being of individuals receiving human services.

Licensing measurement and monitoring systems can present various challenges and considerations, including privacy concerns, data security, resource constraints, and the need for ongoing system updates and maintenance. Addressing these challenges is crucial to ensure the effective implementation and operation of these systems. Additionally, collaboration between regulatory agencies, stakeholders, and technology providers is essential to overcome these challenges and maximize the benefits of these systems.

- **Privacy concerns:** Measurement and monitoring systems often involve the collection and analysis of sensitive data, such as personal information or proprietary business data. It is important to establish robust privacy

policies and legal frameworks to protect individuals' privacy rights and ensure compliance with relevant data protection regulations. Implementing anonymization techniques, data minimization principles, and obtaining appropriate consent can help mitigate privacy concerns.

- **Data security:** The storage, transmission, and analysis of measurement and monitoring data require robust security measures to prevent unauthorized access, data breaches, or cyber-attacks. Encryption, access controls, regular security audits, and adherence to industry best practices can help safeguard the data and maintain its integrity and confidentiality.
- **Resource constraints:** Licensing measurement and monitoring systems can pose financial and logistical challenges, particularly for smaller organizations or developing countries with limited resources. These systems may require substantial investments in infrastructure, equipment, and skilled personnel. Adequate funding mechanisms, public-private partnerships, and capacity-building initiatives can help address resource constraints and ensure broader access to these systems.
- **Ongoing system updates and maintenance:** Measurement and monitoring systems must be regularly updated to keep pace with evolving technologies, regulatory requirements, and scientific advancements. This necessitates ongoing maintenance, software updates, calibration, and quality control procedures. Collaboration between regulatory agencies, technology providers, and stakeholders is crucial to establish effective mechanisms for system maintenance, ensuring that the systems remain accurate, reliable, and up-to-date.
- **Collaboration between regulatory agencies, stakeholders, and technology providers:** Overcoming the challenges associated with licensing measurement and monitoring

systems requires a collaborative approach. Regulatory agencies should engage in constructive dialogues with stakeholders, including industry representatives, environmental organizations, and community groups. Collaboration can help address concerns, establish common standards, and promote transparency and accountability. Technology providers can contribute by developing user-friendly and interoperable systems that meet regulatory requirements while minimizing the burden on end-users.

Collaboration among regulatory agencies, stakeholders, and technology providers is critical to ensure the successful implementation of measurement and monitoring systems. By working together, these entities can develop robust policies, address privacy concerns, enhance data security, allocate necessary resources, and establish mechanisms for ongoing system updates and maintenance. This collaborative approach will maximize the effectiveness of these systems in monitoring and safeguarding various aspects of public health, environmental quality, and regulatory compliance.

Licensing measurement and monitoring systems play a crucial role in human services regulatory administration by ensuring compliance, enhancing service quality, and protecting individuals receiving care. Integrating regulatory science principles into licensing processes further strengthens these benefits.

One significant aspect of licensing measurement and monitoring systems is their ability to promote compliance. These systems provide a standardized framework for evaluating and assessing the compliance of service providers with established regulations and standards. By implementing these systems, regulatory authorities can systematically track and measure compliance levels, identify areas of non-compliance, and take appropriate

actions to rectify any deficiencies. This helps maintain a high level of accountability among service providers, ensuring they adhere to the required standards and regulations.

Moreover, integrating regulatory science principles into licensing processes brings several advantages. Regulatory science applies scientific knowledge and methodologies to inform regulatory decision-making. By incorporating these principles into licensing, regulators can leverage evidence-based approaches to establish standards, design measurement tools, and set performance benchmarks. This approach promotes objectivity, transparency, and consistency in the licensing process, ensuring that decisions are based on sound scientific evidence rather than subjective judgment.

Another key benefit is the potential for improved service quality. Licensing measurement and monitoring systems enable regulators to gather comprehensive data on service providers' performance, outcomes, and service quality indicators. This information allows for a thorough assessment of service delivery, identifying strengths and weaknesses in the system. By analyzing this data, regulators can provide feedback, guidance, and support to service

providers, fostering continuous improvement in service quality. This leads to better outcomes for individuals receiving care and enhances overall service provision within the human services sector.

Furthermore, licensing measurement and monitoring systems are instrumental in protecting the well-being of individuals receiving care. These systems help identify potential risks, such as violations of safety protocols or instances of abuse or neglect. By closely monitoring service providers, regulators can swiftly respond to any issues, take necessary corrective actions, and ensure the safety and well-being of vulnerable populations. Regular monitoring also acts as a deterrent, encouraging service providers to maintain high standards and comply with regulations to avoid penalties or sanctions.

Looking ahead, the field of regulatory science and measurement and monitoring systems is continually evolving. Advances in technology, data analytics, and artificial intelligence present opportunities for further advancements in these systems. For example, the integration of real-time data collection and analysis can enhance the effectiveness and efficiency of monitoring processes. Predictive analytics and risk assessment models can help regulators proactively identify potential areas of concern and allocate resources accordingly. Additionally, the incorporation of feedback from individuals receiving care and other stakeholders can further refine measurement systems, ensuring they capture the most relevant and meaningful indicators of service quality.

In conclusion, licensing measurement and monitoring systems are vital components of human services regulatory administration. By integrating regulatory science principles, these systems promote compliance, improve service quality, and protect individuals receiving care. As regulatory science continues to evolve, the potential for further advancements in measurement and monitoring systems is promising, enabling regulators to better fulfill their mandate of safeguarding the well-being of vulnerable populations.

1 Richard Fiene PhD, Research Psychologist/Regulatory Scientist, Research Institute for Key Indicators; Emeritus Professor of Psychology, Prevention Research Center, Penn State University; Senior Research Consultant, National Association for Regulatory Administration, rfiene@rikinstitute.com

Research Institute for Key Indicators: <https://rikinstitute.com>

Retrieved from <https://encyclopedia.pub/entry/history/show/130069>

Development of a Regulatory Compliance Scale

Subjects: **Psychology**

Contributor: Richard Fiene

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

regulatory science

regulatory compliance

licensing measurement

human services regulatory administration

Development of a Regulatory Compliance Scale

Richard Fiene PhD

Penn State's Edna Bennett Pierce Prevention Research Center

April 2025

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

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

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

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

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

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

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

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

Introduction to the Regulatory Compliance Scale

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

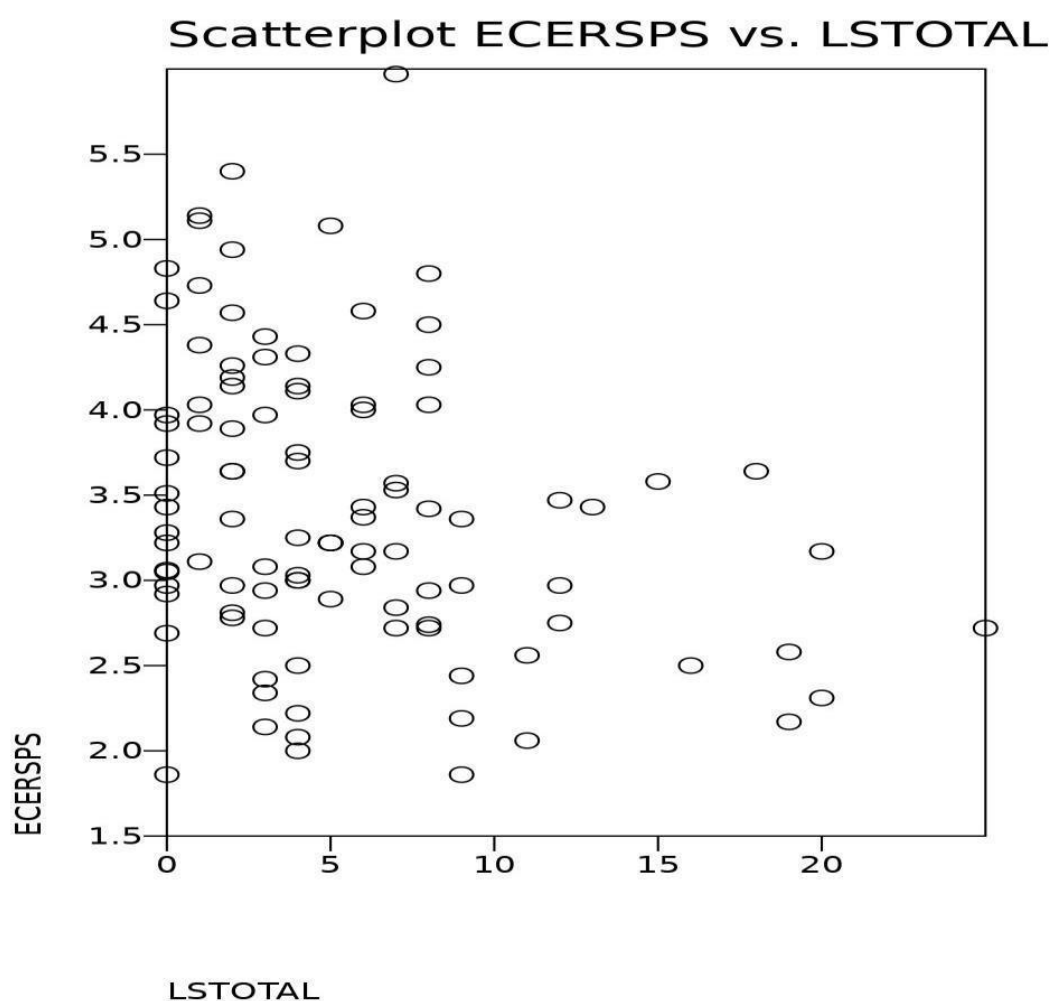
The theory of regulatory compliance has another application when it comes to regulatory compliance measurement in helping to move the licensing field from a nominal based measurement strategy to one of ordinal based measurement. The new measurement strategy is the Regulatory Compliance Scale (RCS) and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

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

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

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

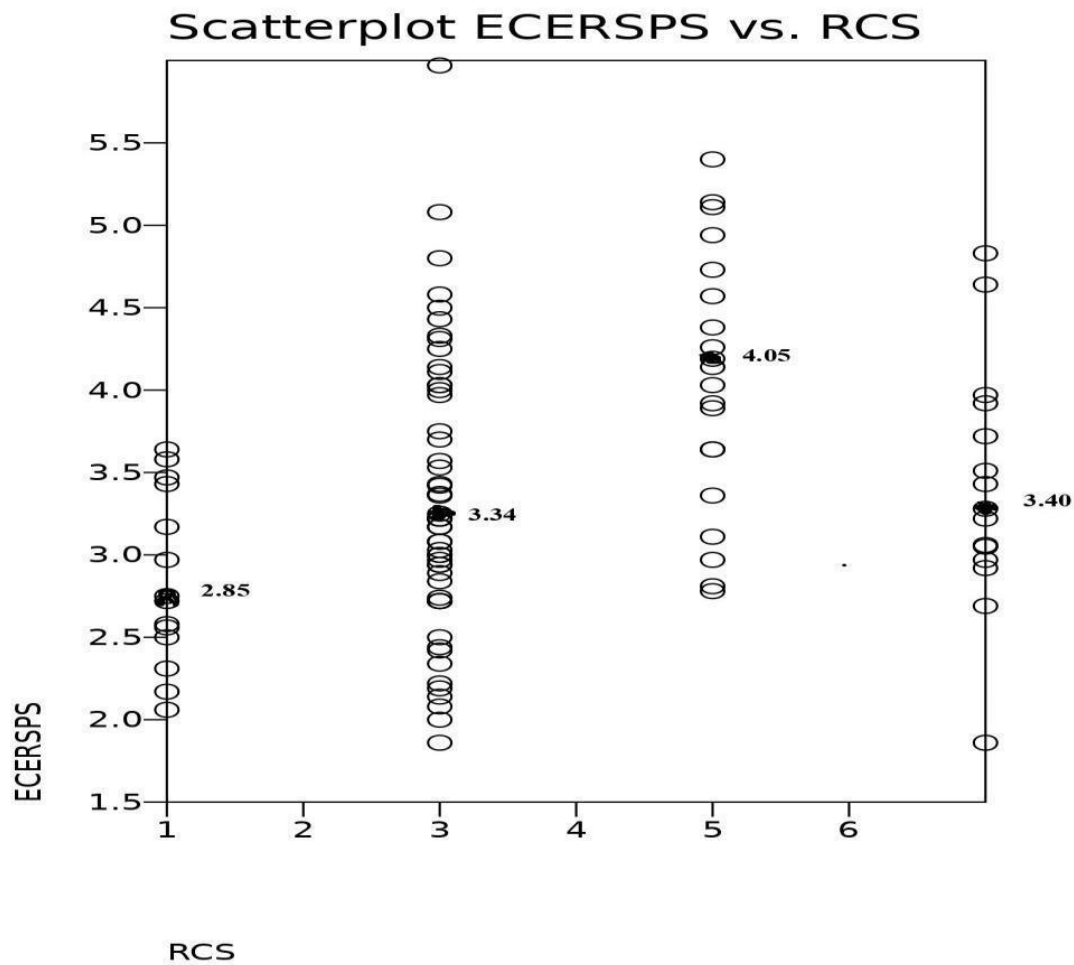
The first figure shows the actual individual violation data of the programs compared to their corresponding ECERS scores. There is not a significant relationship between the two as depicted in the graphic.



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

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

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



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

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

Regulatory Compliance Scale Studies

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

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

Methodology

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

RCS Models used for analyses

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

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

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

Results

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

RCS Model Results compared to Quality Scales

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

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

*Statistically significant .05 level;

**Statistically significant .01 level.

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

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

RCS Fibonacci Models

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

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

ANOVAs Comparing the RCS Models with Program Quality

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

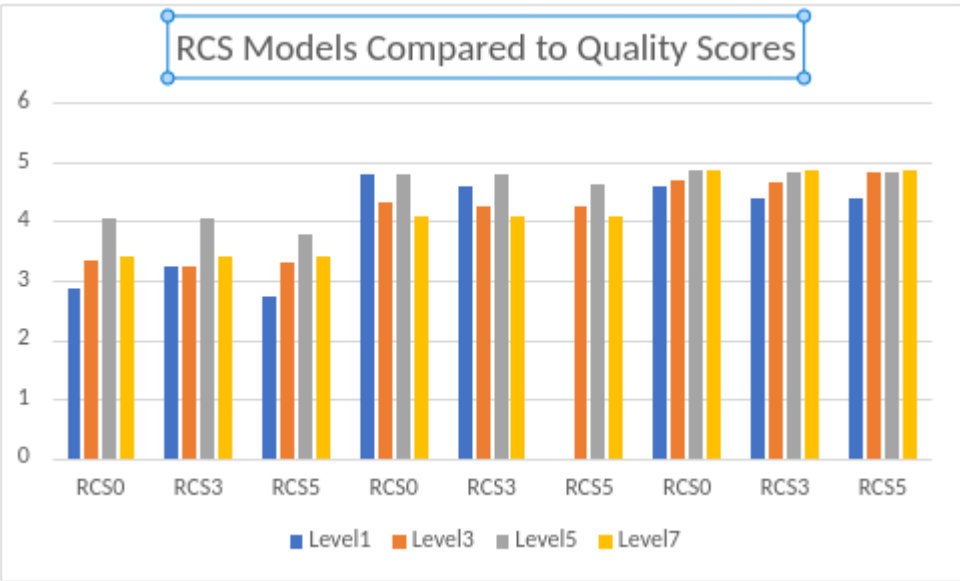
Insights

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

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

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

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



Additional Analyses Comparing the 11 Studies

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

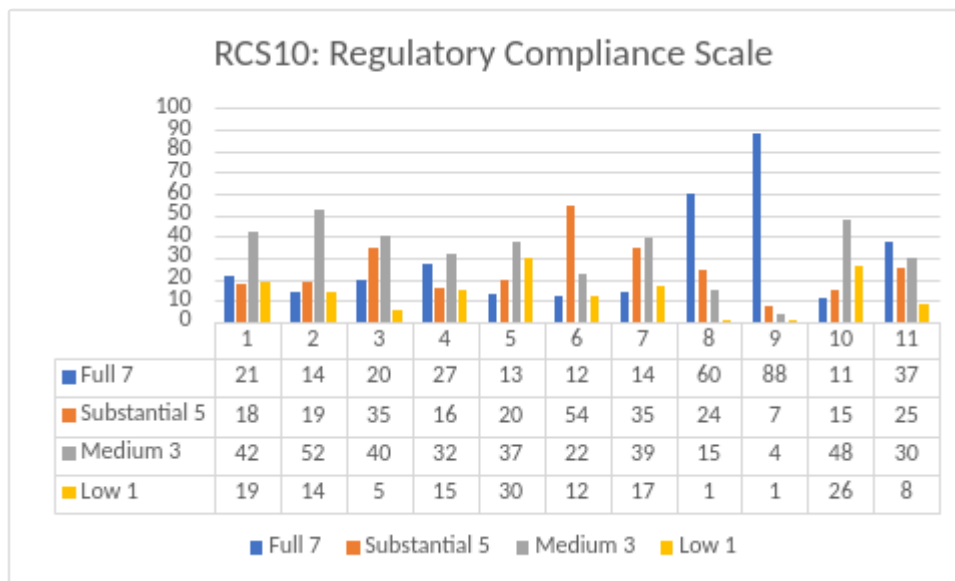
The RCS was based upon the following rubric: Full Regulatory Compliance (100%) or no violations = 7; Substantial Regulatory Compliance (99-98) or 1-2 violations = 5; Medium Regulatory Compliance (97-90) or 3-10 violations =3; and Low Regulatory Compliance (89 or less) or 11 or more violations = 1.

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

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

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

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



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 mitigate the severe skewness in the regulatory compliance data distribution. The data distribution does not approximate a normally distributed curve which is the case with the program quality data distribution.

Regulatory Compliance x Program Quality Grid Model

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

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

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

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

Regulatory Compliance Scoring System and Scale (RC3S)

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

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

Regulatory Compliance Scaling for Decision Making

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

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

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

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

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

Data Level Decision_____

100-98% Full/Substantial License

97-90% Mid-Range Provisional

89% or less Low No-License

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

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

Conclusion

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

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

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

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

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

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

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

Regulatory Compliance Scales and Program Monitoring Systems

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

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

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

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

Regulatory Compliance Scales:

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

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

Instrument Based Program Monitoring (IBPM):

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

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

Differential Monitoring (DM):

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

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

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

Integrative Monitoring Systems (IMS):

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

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

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

Here's a simplified analogy to illustrate the relationships:

Think of regulations as traffic rules.

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

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

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

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

Relationships:

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

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

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

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

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

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

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

Here are some additional points to consider:

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

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

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

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

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

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

References:

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

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

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

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

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

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

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

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

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

Retrieved from <https://encyclopedia.pub/entry/history/show/130024>

Uncertainty-Certainty Matrix Logic Model and Algorithms (UCMLMA)

Subjects: **Psychology**

Contributor: Richard Fiene

Introduces the Uncertainty-Certainty Matrix Logic Model and Algorithms for making human services licensing decisions and avoiding inspection bias and reliability/validity errors.

regulatory compliance

licensing measurement

differential monitoring

logic models

human services licensing

decision making models

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:

Table 1: Uncertainty-Certainty Matrix (UCM) Logic Model

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

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

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):

$$FC = ((A)(D)) - ((B)(C)) / \text{sqrt}(WXYZ)$$

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:

$$FC^* = ((A)(D)) - ((B^3)(C)) / \text{sqrt}(WXYZ)$$

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	<i>TP = True Positive</i>	<i>FP = False Positive</i>
Standard	NC = Non Compliance	<i>FN = False Negative</i>	<i>TN = True Negative</i>

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 $((TP \times TN) - (FN \times FP)) / \sqrt{((TP + FN)(FP + TN)(TP + FP)(FN + TN))}$ 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: $((TP \times TN) - (FN \times FP)) / \sqrt{((TP + FN)(FP + TN)(TP + FP)(FN + TN))}$ 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	<i>TP = True Positive</i>	<i>FP = False Positive</i>
Regulation	NC = Non Compliance	<i>FN = False Negative</i>	<i>TN = True Negative</i>

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: $((TP \times TN) - (FN \times FP) / \sqrt{((TP + FN)(FP + TN)(TP + FP)(FN + TN))})$ 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 $((TP \times TN) - (FN \times FP) / \sqrt{((TP + FN)(FP + TN)(TP + FP)(FN + TN))})$ 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.

Retrieved from <https://encyclopedia.pub/entry/history/show/130320>

A Proposal for A Spatial Acquisition Device

Richard Fiene

Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University

Correspondence: rfiene@rikoinstitute.com

Abstract: This short paper proposes a Spatial Acquisition Device building off of Jean Piaget's theory of cognitive development in how children construct their conceptual framework of spatial dimensions. The proposal demonstrates how children in Piaget's Sensori-motor stage experience reality in three dimensions and then internalize those experiences within their cognitive framework. Developmental psychology implications are addressed as well as broader implications that may extend into other scientific areas in how our notions of space are constructed.

Keywords: Cognitive Development; Developmental Psychology; Young Children; Acquisition Devices.

Introduction

Jean Piaget (1) has influenced cognitive theory with how children develop their concepts of space and time. The theory delineates a four-stage process of moving from sensori-motor to pre-operations to concrete operations and finally to formal operational thought. The developmental process is one of increasing complexity as the child internalizes and builds upon their concepts of space and time. It is almost as if they have a Spatial Acquisition Device, similar to Noam Chomsky's Language Acquisition Device (2).

One of Piaget's most important discoveries is the concept of object permanence in which the child begins to understand that an object exists even when it cannot be viewed. Prior to object permanence when an object is removed from a child's sight, it no longer exists. This acquisition of object permanence occurs in the first two years of life during a child's sensori-motor stage of cognitive development. The child develops the sense of being as having a permanent existence physically. So just as the child has learned about three dimensions physically through movement in the first two-years of life (being stationary, to sitting up, to crawling, and eventually to walking), the child begins a journey of internalizing how they learn about dimensionality over the next three stages of Piaget's theory in moving from one dimension to two dimensions and finally to three- dimensions.

Piaget invented a very unique experiment to test for this acquisition called “conservation experiments” in which he devised experiments for one-dimension (number), two-dimensions (area), and three- dimensions (volume) (see the following table: Developing the Spatial Acquisition Device) in the next Section: The Proposal which will clearly delineate how the above will be reformatted. This is a new way of thinking about the theory. In the past, criticisms have been made in the actual time frame of when various stages appeared, but the sequencing of number to area to volume have never been challenged.

The Proposal

This proposal extends Piaget’s theory of cognitive development into a new conceptual framework in attempting to build in dimensionality into his conservation experiments. His object permanence experiments deal with establishing the notion of an object being permanent to the child, even when it is no longer present. This is a big step because prior to this step, out of mind was no longer existing.

Developing the Spatial Acquisition Device

<i>Dimensionality</i>	<i>Conservation Experiment</i>	<i>Stage of Piaget’s Theory</i>
0 Dimension	Object Permanence	Sensori-Motor
1 Dimension	Number	Late Pre-Operational
2 Dimensions	Area	Concrete Operational
3 Dimensions	Volume	Concrete Operational

This acquisition is invariant, it may occur at different ages for children but they are not going to go from object permanence to area before going to number in conservation experiments, for example. Recent research has demonstrated that Piaget’s stages may occur a bit earlier than suggested by the theory, but the invariance of spatial dimensionality has not been challenged. This is an important discovery since it could lead us to the Spatial Acquisition Device, which has implications in how we interpret the world. Are our brains pre- wired to interpret the world within three-dimensions and if so what could be the next logical step in understanding the relationship between space and time.

Discussion

A new proposed Spatial Acquisition Device has been suggested building off of Jean Piaget’s theory of cognitive development utilizing similar thinking similar to Chomsky’s Language Acquisition Device. There are discussions within the broader scientific arena, in particular physics and quantum mechanics, where the notion of a spatial acquisition device could be a handy tool for understanding how we acquire and construct our particular concept of spacetime. This could be an interesting area of study to determine if this sequencing and construction is invariant and exploring epistemologically the logical conclusions related to spacetime.

References

1. Jean Piaget (1950), *The Construction of Reality in the Young Child*. Basic Books, London, England.
2. Noam Chomsky (1977). *On Language*. The Free Press, New York, New York.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Importance of the Theory of Regulatory Compliance

Richard Fiene

Penn State Prevention Research Center

The purpose of this article is to present the theory of regulatory compliance to a larger management audience. The theory has grown out of research done in the human services but it has not permeated the other industries where the results could be very pertinent. A series of key concepts are presented with a listing of specific elements which would have an impact of a particular industry utilizing the theory of regulatory compliance.

Introduction

Regulatory compliance refers to the process by which individuals, organizations, or entities adhere to and fulfill the requirements set forth by relevant laws, regulations, and industry standards. It involves ensuring that policies, procedures, and practices align with the specific legal and regulatory frameworks applicable to a particular industry or jurisdiction.

Compliance involves actively identifying and understanding the relevant regulations, establishing internal controls and processes to meet those requirements, and consistently monitoring and reviewing operations to ensure ongoing adherence. It encompasses various aspects, such as legal, financial, operational, and ethical considerations, and aims to ensure that organizations operate within the boundaries of the law, maintain ethical standards, and fulfill

their responsibilities to stakeholders, customers, and the public.

The theory of regulatory compliance provides a framework for understanding the underlying principles and concepts that guide the compliance process. It encompasses several key elements that shape the approach to achieving and maintaining compliance. Here is an overview of the theory of regulatory compliance:

1. **Legal and Regulatory Environment:** The theory recognizes that regulatory compliance is rooted in the legal and regulatory landscape. It acknowledges the importance of identifying and understanding applicable laws, regulations, and standards that govern an industry or jurisdiction.
2. **Risk Management:** The theory emphasizes the proactive identification, assessment, and management of risks associated with non-compliance. It highlights the need to establish robust risk management processes to mitigate legal, financial, operational, and reputational risks.
3. **Policies and Procedures:** Effective compliance requires the development and implementation of comprehensive policies and procedures. The theory underscores the significance of clear, well-documented, and communicated policies that guide employees in adhering to regulatory requirements.
4. **Internal Controls:** The theory emphasizes the establishment of internal controls to ensure compliance. This involves designing and implementing systems, processes, and checks that monitor and mitigate risks, detect and prevent non-compliance, and promote accountability.
5. **Training and Awareness:** Recognizing the role of individuals in compliance, the theory highlights the importance of training programs and awareness initiatives. It emphasizes educating

employees about applicable regulations, ethical standards, and the organization's compliance obligations.

6. **Monitoring and Auditing:** The theory acknowledges the need for ongoing monitoring and auditing to assess compliance effectiveness. Regular internal audits, reviews, and assessments help identify gaps, weaknesses, and areas for improvement, ensuring continuous compliance efforts.
7. **Reporting and Documentation:** The theory stresses the significance of accurate and timely reporting of compliance activities. It underscores the need to maintain proper documentation, records, and evidence of compliance processes, actions taken, and outcomes achieved.
8. **Compliance Culture:** The theory recognizes that compliance is not solely a set of rules and processes but also a cultural mindset. It highlights the importance of fostering a culture of compliance within an organization, where integrity, ethics, and adherence to regulations are valued and embedded in the organizational DNA.
9. **Accountability and Enforcement:** The theory acknowledges that compliance requires accountability for non-compliance. It recognizes the role of regulatory bodies, internal enforcement mechanisms, and disciplinary actions in promoting compliance and deterring violations.
10. **Continuous Improvement:** Finally, the theory emphasizes the need for continuous improvement in compliance efforts. It encourages organizations to learn from past experiences, adapt to evolving regulations, embrace emerging best practices, and strive for excellence in their compliance initiatives.

By understanding and applying the theory of regulatory compliance, organizations can

establish a solid foundation for effective compliance management, minimize risks, and uphold legal and ethical standards in their operations.

Ensuring Legal and Ethical Practices

Compliance with laws and regulations is a fundamental aspect of the theory of regulatory compliance. It recognizes that adherence to legal and regulatory requirements is crucial for organizations to operate within the boundaries set by governing bodies and to fulfill their obligations to stakeholders. Here are key points related to compliance with laws and regulations in the context of the theory of regulatory compliance:

11. **Understanding Applicable Laws:** The theory emphasizes the importance of identifying and comprehending the specific laws and regulations that pertain to an organization's industry, jurisdiction, and operational activities. This involves staying updated with changes in regulations and interpreting their implications for the organization.
12. **Regulatory Research and Analysis:** Organizations need to conduct thorough research and analysis to determine how laws and regulations apply to their operations. This includes examining regulatory frameworks, guidance documents, legal precedents, and industry-specific requirements.
13. **Compliance Obligations:** The theory recognizes that compliance obligations vary based on the nature of the organization's activities. It stresses the need to determine the specific requirements, obligations, and standards that the organization must meet to ensure legal and regulatory compliance.
14. **Compliance Program Development:** To achieve compliance, the theory highlights the importance of developing a comprehensive compliance program

tailored to the organization's needs. This involves establishing policies, procedures, and controls that align with legal and regulatory requirements.

15. **Regulatory Reporting and Filings:** Compliance entails fulfilling reporting obligations to regulatory authorities. The theory emphasizes the significance of timely and accurate reporting, including financial statements, disclosures, permits, licenses, certifications, and other regulatory filings.
16. **Compliance Monitoring and Auditing:** The theory underscores the need for ongoing monitoring and auditing of compliance efforts. Regular reviews help identify potential compliance gaps, assess the effectiveness of controls, and ensure corrective actions are taken to address non-compliance.
17. **Compliance Documentation:** Documentation plays a critical role in compliance. The theory highlights the importance of maintaining accurate and comprehensive records of compliance activities, including policies, procedures, training materials, audit reports, incident reports, and evidence of compliance.
18. **Compliance Risk Assessment:** Organizations should conduct compliance risk assessments to identify and evaluate potential risks associated with non-compliance. This allows for the implementation of risk mitigation strategies, such as internal controls, training programs, and monitoring systems.
19. **Enforcement and Consequences:** The theory acknowledges that non-compliance can lead to legal and financial consequences. It emphasizes the need for organizations to understand the potential penalties, fines, sanctions, and reputational damage that can result from violations of laws and regulations.
20. **Regulatory Engagement and Communication:** Organizations should

actively engage with regulatory authorities and maintain open lines of communication. The theory emphasizes the importance of understanding regulatory expectations, seeking guidance when needed, and participating in industry consultations.

By emphasizing compliance with laws and regulations, the theory of regulatory compliance aims to ensure that organizations operate within legal boundaries, mitigate risks, protect stakeholders, and maintain a strong ethical foundation in their operations.

Protection of Consumers and Public Interest

Protection of consumers and the public interest is a fundamental objective of regulatory compliance. Regulatory compliance refers to the adherence of individuals, organizations, or businesses to laws, regulations, and guidelines set forth by governing bodies or regulatory authorities. It aims to ensure that entities operate in a manner that safeguards the interests of consumers and the general public.

The theory behind regulatory compliance is rooted in the belief that certain industries or activities require oversight and regulation to prevent harm, ensure fair competition, and maintain public trust. By establishing rules and standards, regulatory bodies seek to create a level playing field, promote transparency, and protect the well-being of consumers.

Key principles and considerations associated with regulatory compliance for the protection of consumers and public interest include:

21. **Consumer Protection:** Regulatory compliance frameworks typically include provisions to safeguard consumers from fraudulent, deceptive, or unfair practices. This involves regulations related to product safety,

labeling, advertising, pricing, warranties, and consumer rights.

22. **Public Health and Safety:** Compliance regulations often address public health and safety concerns. For instance, in the pharmaceutical industry, compliance with drug safety regulations ensures that medications meet quality standards and do not pose unreasonable risks to patients.
23. **Market Integrity:** Regulatory compliance helps maintain the integrity of markets by prohibiting anti-competitive behavior, ensuring fair trading practices, and preventing market manipulation or insider trading. These regulations promote fair competition and protect consumers from monopolistic practices.
24. **Data Protection and Privacy:** With the increasing prevalence of data-driven technologies, regulatory compliance frameworks emphasize the protection of personal information and privacy rights. Regulations like the European Union's General Data Protection Regulation (GDPR) aim to safeguard consumer data and establish guidelines for its lawful collection, storage, and use.
25. **Financial Stability:** Regulatory compliance plays a crucial role in the financial sector to prevent fraud, money laundering, and unethical practices that can destabilize markets or harm consumers. Regulations impose standards for capital adequacy, risk management, disclosure, and consumer financial protection.
26. **Ethical Considerations:** Compliance regulations often incorporate ethical considerations to ensure responsible and ethical behavior by individuals and organizations. This may involve guidelines on corporate governance, social responsibility, environmental sustainability, or labor practices.

To ensure effective regulatory compliance, regulatory bodies conduct inspections, audits, and enforcement actions. Non-compliance can result in penalties, fines, or legal actions against the offending parties. Moreover, compliance management systems, internal controls, and self-regulatory mechanisms are employed by organizations to proactively adhere to regulatory requirements and promote a culture of compliance.

Overall, the theory of regulatory compliance revolves around the idea that by setting and enforcing rules, regulators can protect consumers, preserve public interest, and maintain the stability and fairness of various sectors in society.

Financial Stability and Risk Management

Financial stability and risk management are critical components of regulatory compliance. The theory of regulatory compliance emphasizes the importance of establishing and enforcing regulations to ensure the stability and integrity of financial systems, protect consumers, and mitigate systemic risks.

Here are some key aspects of the theory of regulatory compliance related to financial stability and risk management:

27. **Prudential Regulation:** Prudential regulation focuses on ensuring the soundness and stability of financial institutions, such as banks, insurance companies, and investment firms. Regulatory compliance frameworks impose requirements related to capital adequacy, risk management, liquidity, and asset quality to prevent excessive risk-taking and protect the financial system from disruptions.
28. **Systemic Risk Mitigation:** Regulatory compliance measures aim to identify and mitigate systemic risks that can have widespread adverse effects on the

financial system. This includes regulations on risk concentration, interconnectedness, and exposure limits to prevent the domino effect of failures and contagion across institutions.

29. Risk Assessment and Monitoring: Regulatory compliance frameworks often require financial institutions to conduct thorough risk assessments and implement robust risk management practices. This involves identifying, measuring, and monitoring various types of risks, including credit risk, market risk, liquidity risk, and operational risk. Compliance regulations may prescribe specific methodologies, reporting requirements, and stress testing to ensure that risks are adequately identified and managed.

30. Transparency and Disclosure: Regulatory compliance promotes transparency in financial markets by requiring financial institutions to provide accurate and timely disclosure of relevant information to investors, regulators, and the public. This includes financial reporting, disclosures of risk exposures, and information about the institution's financial health. Transparent reporting helps stakeholders make informed decisions, enhances market efficiency, and fosters trust in the financial system.

31. Consumer Financial Protection: Regulatory compliance frameworks incorporate measures to protect consumers in financial transactions. This includes regulations on fair lending practices, disclosure requirements for financial products and services, and regulations against abusive or predatory practices. These regulations aim to ensure that consumers are treated fairly, have access to transparent information, and are protected from fraudulent or deceptive practices.

32. Regulatory Oversight and Enforcement: Regulatory compliance is reinforced by regulatory bodies that oversee financial institutions, enforce compliance, and impose penalties for non-compliance. These regulatory authorities monitor institutions' compliance with regulations, conduct audits and examinations, and take enforcement actions when violations are identified. Such oversight ensures accountability and promotes a culture of compliance within the financial industry.

By adhering to regulatory compliance requirements, financial institutions are expected to minimize risks, enhance stability, and maintain the confidence of investors and the public. Compliance management systems, internal controls, and risk management frameworks are utilized by financial institutions to meet regulatory obligations and proactively manage risks.

Overall, the theory of regulatory compliance underscores the role of regulations in promoting financial stability, mitigating risks, protecting consumers, and maintaining the integrity of financial systems. Compliance with these regulations helps build a resilient financial sector that can withstand shocks and contribute to overall economic stability.

Preserving Competitive Market Environment

Preserving a competitive market environment is essential for fostering innovation, encouraging efficiency, and benefiting consumers. The theory of regulatory compliance is closely linked to this objective, as it involves establishing and enforcing rules and regulations that promote fair competition and prevent anti-competitive practices.

The theory of regulatory compliance is based on the idea that regulatory frameworks can help create a level playing field for all market participants. By setting clear rules and standards,

regulators aim to ensure that businesses operate within the bounds of fair competition. Compliance with these regulations helps prevent monopolistic behavior, collusion, price-fixing, and other practices that could harm competition.

Here are a few key principles related to preserving a competitive market environment and the theory of regulatory compliance:

33. **Anti-Trust Laws:** Anti-trust laws are designed to promote competition by preventing the abuse of market power. They prohibit practices such as monopolies, cartels, price-fixing, and mergers that may substantially lessen competition. Regulators enforce these laws to preserve a competitive landscape and protect consumer interests.
34. **Market Entry and Exit:** Regulatory frameworks should facilitate the entry of new businesses into the market while allowing existing ones to exit if they are unable to compete effectively. Barriers to entry, such as excessive licensing requirements or unfair regulations, can hinder competition. Regulatory compliance should aim to reduce these barriers and ensure fair access for all participants.
35. **Consumer Protection:** A competitive market environment should prioritize consumer welfare. Regulatory compliance plays a crucial role in safeguarding consumer interests by ensuring transparency, fair pricing, quality standards, and adequate information disclosure. Consumer protection laws and regulations address issues such as misleading advertising, product safety, and fair dispute resolution mechanisms.
36. **Enforcement and Monitoring:** Regulatory agencies are responsible for enforcing compliance with regulations. They monitor market activities, investigate potential violations, and take

appropriate enforcement actions when necessary. Effective enforcement requires sufficient resources, expertise, and collaboration among regulators, ensuring a level playing field for all participants.

37. **International Cooperation:** In a globalized economy, preserving a competitive market environment requires international cooperation. Collaboration between regulatory authorities across jurisdictions can help address cross-border anti-competitive practices, harmonize regulatory standards, and promote fair competition in the global marketplace.

Overall, the theory of regulatory compliance supports the notion that well-designed and effectively enforced regulations can foster a competitive market environment. By promoting fair competition, preventing anti-competitive practices, and protecting consumer interests, regulatory compliance contributes to a healthy and vibrant marketplace.

Establishing Trust and Credibility

Establishing trust and credibility is crucial for regulatory compliance efforts. The theory of regulatory compliance recognizes that trust is essential in fostering cooperation between regulatory authorities, businesses, and other stakeholders. Trust is built when regulations are transparent, consistently enforced, and perceived as fair and unbiased.

Here are some key aspects of establishing trust and credibility in the context of regulatory compliance:

38. **Transparency:** Transparency is a fundamental principle in regulatory compliance. Regulations and their enforcement processes should be clearly communicated and accessible to all stakeholders. Openness helps build trust

by ensuring that the rules are known and understood by businesses and individuals, reducing uncertainty and promoting voluntary compliance.

39. **Consistency:** Consistency in applying regulations is critical for building trust. Regulators should strive to enforce regulations uniformly and without favoritism or discrimination. Consistent enforcement establishes a level playing field, fostering trust among market participants who know that everyone is subject to the same rules.
40. **Accountability:** Regulatory authorities should be accountable for their actions. This includes being transparent about decision-making processes, justifying regulatory actions, and providing avenues for recourse and appeal. Accountability mechanisms help prevent abuse of regulatory power and build trust by demonstrating fairness and impartiality.
41. **Collaboration and Engagement:** Regulatory compliance efforts benefit from collaboration and engagement with various stakeholders. This includes businesses, industry associations, consumer groups, and experts. Involving stakeholders in the regulatory process helps ensure that regulations are practical, effective, and well-understood. Collaboration also enhances trust by incorporating diverse perspectives and building consensus.
42. **Risk-Based Approach:** A risk-based approach to regulation can contribute to trust and credibility. It involves assessing risks, prioritizing enforcement efforts based on the potential harm to the public or the market, and proportionately allocating regulatory resources. This approach demonstrates that regulatory actions are driven by objective evaluations and the need to address significant risks, enhancing trust in the regulatory system.

43. **Continuous Improvement:** Regulatory compliance should be a dynamic and evolving process. Regular evaluation and improvement of regulations and enforcement mechanisms are essential for maintaining trust and credibility. Regulators should engage in periodic reviews, solicit feedback from stakeholders, and adapt regulations to changing market dynamics and emerging challenges.

44. **Effective Communication:** Clear and effective communication is vital for establishing trust. Regulators should communicate expectations, obligations, and changes in regulations in a timely and accessible manner. Communication channels should be open to addressing queries, providing guidance, and clarifying regulatory requirements, fostering trust by ensuring transparency and promoting compliance.

In summary, trust and credibility are foundational elements of successful regulatory compliance. By promoting transparency, consistency, accountability, collaboration, and effective communication, regulatory authorities can establish a trusted regulatory framework that fosters compliance and cooperation among stakeholders.

Penalties and Consequences of Non-Compliance

Regulatory compliance refers to the act of adhering to laws, regulations, guidelines, and standards set forth by governing bodies or regulatory agencies. Non-compliance occurs when individuals, organizations, or businesses fail to meet these requirements. The penalties and consequences of non-compliance can vary depending on the specific regulations and jurisdictions involved. Here are some common penalties and consequences:

45. **Fines and Monetary Penalties:** Regulatory agencies often have the authority to impose fines and monetary penalties for non-compliance. The amount of the penalty may vary depending on the severity of the violation and the regulatory framework in place. These fines can be substantial and can significantly impact the finances of non-compliant entities.
46. **Legal Proceedings and Lawsuits:** Non-compliance may lead to legal action, including lawsuits filed by affected parties or regulatory bodies. This can result in costly litigation, potential damages, and a tarnished reputation.
47. **License Revocation or Suspension:** Certain industries and professions require licenses or permits to operate legally. Non-compliance can lead to the revocation or suspension of these licenses, effectively shutting down the business or preventing individuals from practicing their profession.
48. **Regulatory Audits and Inspections:** Regulatory agencies may conduct audits and inspections to assess compliance. Non-compliant entities may face increased scrutiny, additional audits, or more frequent inspections, leading to disruption of operations and additional costs.
49. **Reputational Damage:** Non-compliance can harm an organization's reputation, leading to loss of customer trust, decreased sales, and difficulty attracting new customers. Negative publicity and media attention can have long-lasting effects on brand value and perception.
50. **Corrective Actions and Remediation Costs:** In many cases, non-compliant entities are required to take corrective actions to address the violations. This may involve implementing new policies, procedures, or systems, as well as investing in training and education. The

costs associated with these remediation efforts can be significant.

51. **Criminal Charges and Penalties:** In cases of serious non-compliance, intentional violations, or fraudulent activities, criminal charges may be pursued. This can result in fines, imprisonment, or both, depending on the severity of the offense.

The theory of regulatory compliance seeks to understand why individuals or organizations choose to comply or not comply with regulations. Factors influencing compliance behavior include perceived legitimacy of regulations, trust in regulatory agencies, the presence of effective enforcement mechanisms, and the perceived costs and benefits of compliance. The theory emphasizes the importance of clear communication, consistent enforcement, and proportionate penalties to achieve higher compliance rates.

Compliance Programs and Frameworks

Compliance programs and frameworks are designed to help organizations establish and maintain a culture of regulatory compliance. They provide a structured approach to understanding and meeting regulatory requirements, mitigating risks, and promoting ethical behavior. Additionally, compliance programs help organizations detect and address non-compliance issues promptly and effectively.

Here are some common compliance programs and frameworks:

52. **Compliance Management System (CMS):** A CMS is a comprehensive framework that encompasses policies, procedures, processes, and controls to manage compliance within an organization. It includes elements such as risk assessment, compliance training, monitoring and auditing, incident reporting, and corrective action planning.

53. ISO 19600: This international standard provides guidelines for establishing, implementing, evaluating, and improving a compliance management system. It emphasizes a risk-based approach to compliance and provides a framework for organizations to identify, analyze, and address their compliance obligations effectively.
54. COSO Framework: The Committee of Sponsoring Organizations of the Treadway Commission (COSO) developed a framework that focuses on internal controls and risk management. While not specifically geared towards compliance, it provides a solid foundation for managing compliance risks within an organization.
55. Federal Sentencing Guidelines (FSG): The U.S. Federal Sentencing Guidelines provide guidance for organizations on establishing effective compliance programs. They outline specific factors that organizations should consider when developing compliance programs, such as conducting risk assessments, implementing training and communication programs, and monitoring compliance.
56. Principle-Based Approach: The principle-based approach to compliance focuses on establishing a set of core principles and values that guide an organization's compliance efforts. It emphasizes ethical conduct, integrity, and accountability as the foundation for compliance programs. This approach encourages employees to make ethical decisions and act in accordance with the organization's values.

The theory of regulatory compliance explores the factors that influence compliance behavior and the effectiveness of compliance programs. It recognizes that compliance is not solely driven by the fear of penalties but also by factors such as organizational culture, perceived legitimacy of

regulations, and the presence of strong internal controls. The theory suggests that effective compliance programs should:

57. Clearly communicate regulatory requirements and expectations to employees and stakeholders.
58. Foster a culture of compliance by promoting ethical behavior, accountability, and integrity.
59. Provide training and education to employees to enhance their understanding of compliance obligations.
60. Implement monitoring and auditing mechanisms to detect and address non-compliance promptly.
61. Establish strong internal controls and risk management processes to mitigate compliance risks.
62. Encourage reporting of potential compliance issues and provide channels for anonymous reporting.
63. Continuously evaluate and improve the compliance program based on feedback and changes in regulations.

By understanding the theory of regulatory compliance and implementing effective compliance programs, organizations can enhance their ability to meet regulatory requirements, manage risks, and uphold ethical standards.

Role of Technology in Regulatory Compliance and monitoring and reporting tools

Technology plays a crucial role in regulatory compliance by providing tools and systems that help organizations monitor and report their adherence to regulatory requirements. Here are some key ways technology supports regulatory compliance:

64. Automation and Workflow Management: Technology enables the automation of various compliance processes, such as data collection,

analysis, and reporting. Workflow management systems help streamline compliance tasks by providing clear processes and guidelines, ensuring consistent and efficient execution.

65. Data Management and Analysis: Compliance often involves handling large volumes of data. Technology solutions, such as data management systems and analytics tools, facilitate the collection, storage, organization, and analysis of data for compliance purposes. These systems can identify patterns, anomalies, and trends in the data, helping organizations detect and address compliance risks.

66. Monitoring and Surveillance: Technology enables real-time monitoring and surveillance of activities, transactions, and communications to identify potential compliance violations. Advanced monitoring tools use algorithms and machine learning techniques to detect suspicious behavior, fraud, market manipulation, or any non-compliant activities.

67. Reporting and Documentation: Compliance requires accurate and timely reporting to regulatory authorities. Technology offers reporting tools that help automate the creation of regulatory reports, ensuring the required information is captured, organized, and submitted in the appropriate format. These tools often include templates, data mapping capabilities, and integration with existing systems.

68. Audit Trail and Documentation Management: Technology allows organizations to maintain a comprehensive audit trail and documentation of compliance activities. Digital systems enable the secure storage, retrieval, and tracking of compliance-related documents, making it easier to demonstrate compliance during audits or investigations.

69. Risk Assessment and Compliance Monitoring: Technology supports risk assessment processes by providing tools for identifying, assessing, and prioritizing compliance risks. Compliance monitoring tools can continuously track regulatory changes and updates, ensuring organizations stay informed and adapt their compliance programs accordingly.

70. Training and Education: Technology can be utilized to deliver compliance training and educational materials to employees and stakeholders. Online learning platforms, webinars, and interactive modules can provide accessible and engaging compliance training programs, ensuring widespread understanding of regulatory requirements and promoting a culture of compliance.

Overall, technology plays a vital role in enhancing the efficiency, accuracy, and effectiveness of regulatory compliance efforts. By leveraging technology, organizations can better manage compliance requirements, mitigate risks, and ensure adherence to regulations in an increasingly complex regulatory landscape.

Conclusion Recap of the importance of the theory of regulatory compliance

The theory of regulatory compliance is of great importance in various domains, particularly in legal and business contexts. It refers to the set of rules, regulations, and standards that individuals, organizations, and industries must follow to ensure compliance with applicable laws and regulations.

Here are some key points highlighting the importance of the theory of regulatory compliance:

71. Legal Compliance: Regulatory compliance ensures that individuals and organizations adhere to laws and

regulations set forth by governing bodies. This helps maintain law and order in society and promotes fairness, transparency, and accountability.

72. Risk Mitigation: Compliance measures help identify and mitigate potential risks associated with non-compliance. By following regulations, organizations can minimize legal, financial, reputational, and operational risks. Compliance frameworks often include risk assessment and management components, enabling proactive risk mitigation.

73. Consumer Protection: Compliance regulations often aim to protect consumers' rights and interests. Compliance with consumer protection laws ensures fair business practices, prevents fraud, and enhances consumer trust in products and services.

74. Data Privacy and Security: In the digital age, data privacy and security have become crucial concerns. Regulatory compliance frameworks, such as the General Data Protection Regulation (GDPR), enforce strict guidelines for handling personal data. Compliance helps safeguard sensitive information, maintain privacy, and prevent data breaches.

75. Ethical Standards: Compliance extends beyond legal obligations and encompasses ethical standards. It encourages organizations to adopt ethical business practices, such as fair competition, anti-corruption measures, and environmental sustainability. Compliance frameworks often incorporate ethical guidelines to promote responsible conduct.

76. Industry Standards: Many industries have specific regulatory compliance requirements tailored to their unique characteristics and risks. Compliance with industry-specific regulations ensures safety, quality, and

standardization within the sector. Examples include regulations in healthcare, finance, energy, and manufacturing.

77. Reputation and Trust: Compliance with regulations builds a positive reputation for individuals and organizations. It demonstrates commitment to legal and ethical standards, fostering trust among customers, investors, and other stakeholders. A strong reputation for compliance can lead to increased business opportunities and competitive advantage.

78. Legal Consequences: Non-compliance with regulatory requirements can have severe legal consequences, including fines, penalties, sanctions, and legal liabilities. Violations can result in damaged reputation, loss of business licenses, and even criminal charges. Compliance helps organizations avoid legal pitfalls and maintain a good standing with regulatory authorities.

79. Global Business Landscape: With increasing globalization, organizations often need to navigate complex regulatory frameworks across multiple jurisdictions. Understanding and complying with international regulations is essential for expanding businesses, facilitating international trade, and avoiding legal disputes.

80. Continuous Improvement: The theory of regulatory compliance emphasizes the need for continuous improvement. Compliance programs encourage regular monitoring, self-assessment, and adaptability to evolving regulations. This fosters a culture of compliance and enables organizations to stay up to date with changing legal requirements.

In summary, the theory of regulatory compliance plays a vital role in promoting legality, ethical conduct, risk management, and trust in various domains. It ensures adherence to laws, protects

consumers, mitigates risks, and helps organizations thrive in a complex regulatory landscape.

References:

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

Fiene, R. (2022). Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary, *Journal of Regulatory Science: Vol. 10 No. 1 (2022)* <https://doi.org/10.21423/JRS-V10A239>

Regulatory Compliance and Quality Programs: Constraints and Opportunities for Integration

Brian Freer^{a1}, PhD and Richard Fiene^b, PhD ORCID/0000-0001-6095-5085

^aCenter for Science, Technology, Medicine and Society; University of California, Berkeley. Corresponding Author: bjfreer@berkeley.edu

^bEdna Bennett Pierce Prevention Research Center, Pennsylvania State University.

Abstract

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

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

1. Introduction

This paper argues that by examining the costs and benefits of implementing process-based integrated management systems, organizations may gain insight into the potential value of merging regulatory compliance with quality programs. In this paper, we define regulatory compliance as all government requirements, exclusive of accounting, facing an organization and the activities an organization takes to conform to these requirements. Quality programs are defined as customer, international and national standards, and other requirements where an organization is obligated to show conformance. According to ReVelle (2003), a process is “a series of sequentially oriented, repeatable events that have both a beginning and an end, and which result in either a product or a service.” Research by

Carvalho et al. (2015) makes a strong case for integrating multiple standards (e.g., quality, environmental, safety) into a single management system. Their research findings are outlined further below. They define an integrated management system as “a set of interrelated processes that share human resources, information, materials, infrastructure, (sic) financial resources” (Carvalho et al. 2015). However, their research focuses exclusively on international standards, setting aside questions concerning regulatory compliance. Fiene (2019, 2022) has recently made the case that regulatory compliance programs may be enhanced by incorporating measurement and continual improvement—hallmarks of quality initiatives—into regulatory compliance programs. Our focus here is on the challenges and potential efficiencies that organizations may experience through

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

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

Scholars have identified a lack of research on the topic of the relationship between quality programs and regulatory compliance (Doyle 2007; Doyle et al. 2014).

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

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

This alignment of ISO 13485 and 21 CFR 820 is atypical and not yet conventionally

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

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

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

2. Common Problems Encountered by Organizations in Management System Implementation

Organizations inevitably encounter through external audits, executive reviews, and operations, inter-relationships between regulatory compliance and quality programs. However, research suggests that organizations do not generally prioritize investment in the design and implementation of management systems focused on controlling and optimizing this inter-relationship (Doyle 2007). Thus, outside of the realm of operations the relationship between compliance and quality in organizations is more often reacted to in a haphazard manner as opposed to intentionally integrating the two in a management system. As Doyle (2007) has pointed out, one reason for this predicament is that it is inherently difficult for organizations to coordinate legal and supra-legal requirements. Supra-legal refers to binding requirements faced by an organization in addition to government rules. Supra-legal requirements may include national association standards and customer

specifications. Additionally, research has identified other roadblocks that deter organizations from integrating quality programs with regulatory compliance, including limited resources, lack of top management support, and inherent complexity (Doyle et al 2014). Scholars, with the notable exceptions cited above, generally treat the two topics separately. Government agencies and standards registrars, for their part, have historically avoided prescribing structural/organizational requirements concerning documents and their information format in management systems. In part due to this, government and certifying body auditors encounter a myriad of information management schemes in stand-alone and integrated management systems.

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

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

organization, beyond just satisfying regulatory compliance.

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

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

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

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

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

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

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

3. Characterizing Management System Attributes in an Organization

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

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

process oriented with process flowcharts and process maps.

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

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

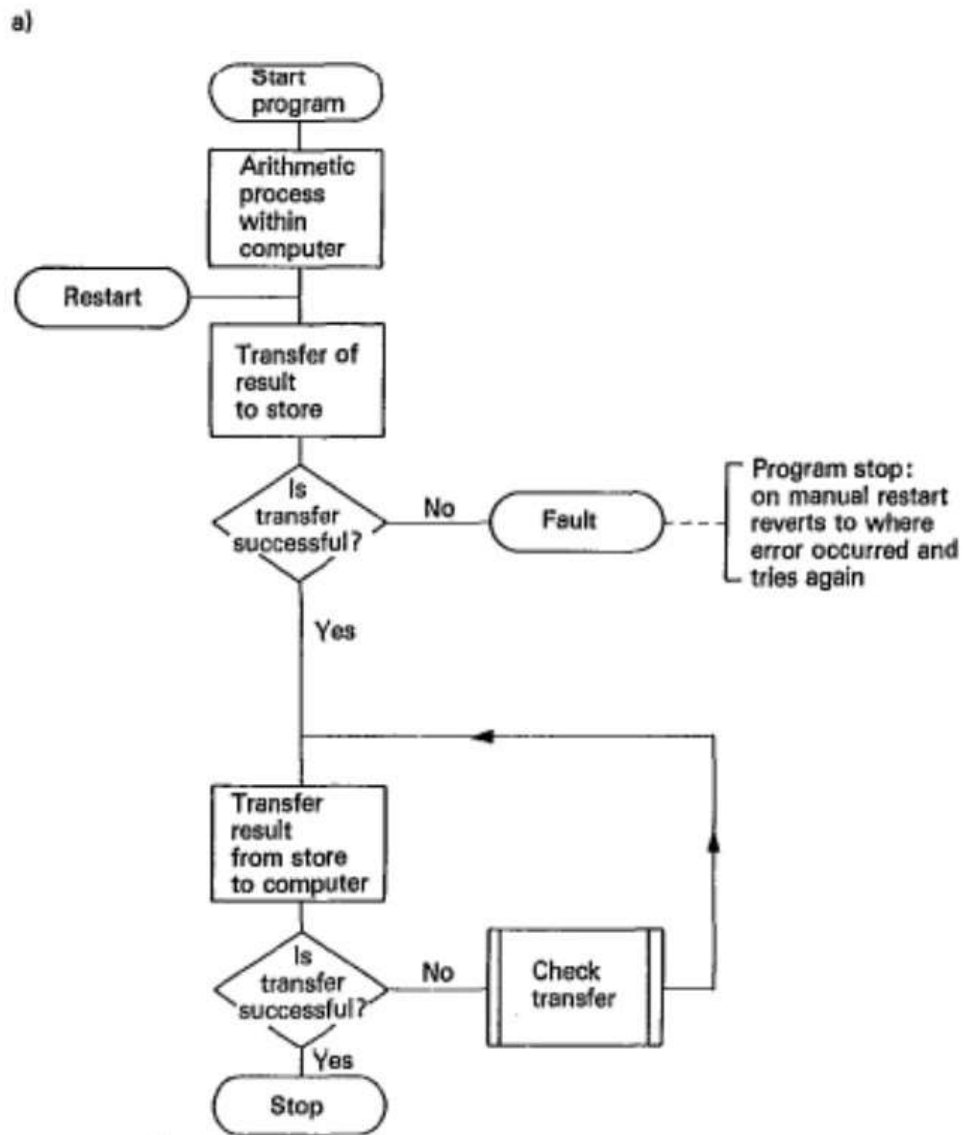


Figure 1: Program Flowchart, from ISO 5087 Annex B

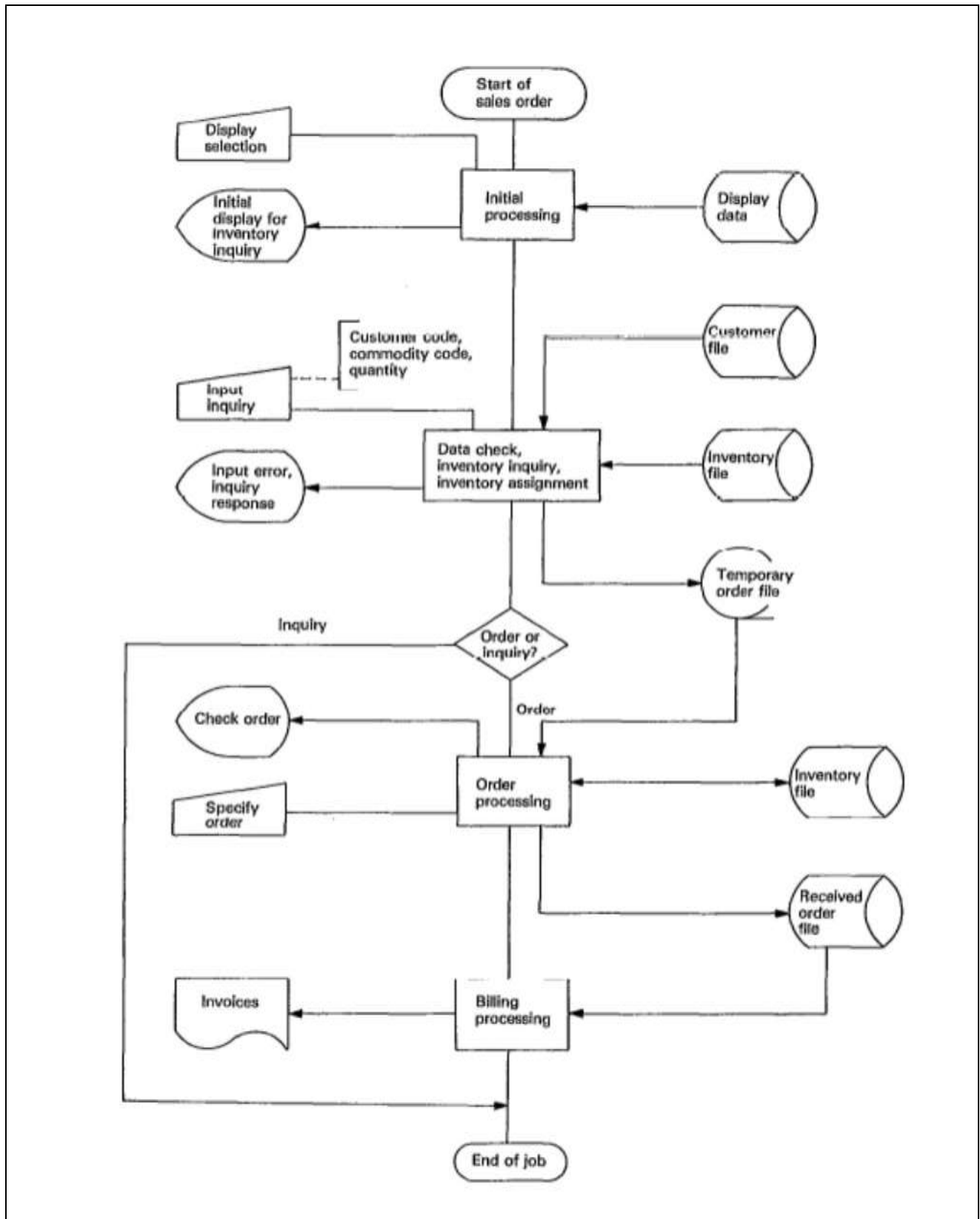


Figure 2: System Flowchart, from ISO 5087, Annex C

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

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

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

To score the overall structure of the management system, assign a label of Elements for systems that follow the sequence and nomenclature of the regulation and/or standard. If the system is based on the organization's own approach, label it independent. For procedures, assign a value of Primarily Narrative for management systems where most of the information in

the procedures and work instructions is in sentences of text without graphical flowcharts. If an organization encounters a management system where the procedures and work instruction are mainly composed of flowcharts and maps, assign a value of Primarily Process.

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

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

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

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

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

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

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

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

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

5. References

- [1] Carvalho, K., Picci, F., Camarini, G., & Chamon (2015). Benefits in the implementation of safety, health, environmental and quality integrated system. *International Journal of Engineering and Technology* 7, no. 4, 355.
- [2] Doyle, E. (2007). Compliance obstacles to competitiveness. *Corporate Governance: The international journal of business in society*, 612-622.
- [3] Doyle, E., McGovern, D., & McCarthy, S. (2014). Compliance–innovation: integrating quality and compliance knowledge and practice. *Total Quality Management & Business Excellence* 25, no. 9-10, 1156-1170.
- [4] Fiene, R. (2022). Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary. *Journal of Regulatory Science* 10, no. 1, 1-7.
- [5] Fiene, R. (2019). A treatise on theory of regulatory compliance, *Journal of Regulatory Science*, 7, no. 1, 1-3.
- [6] International Organization for Standardization. (1985). Information processing – documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts. (ISO Standard No. 5807: 1985).
- [7] Panitz, J. C., Wiener, M., & Amberg, M. (2011) Factors facilitating compliance implementation – Case study results from multinational enterprises, ECIS 2011 Proceedings, 3.

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

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

Richard Fiene, PhD ^{a,*}

^a Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University

Abstract

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

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

Introduction

This commentary on policy will deal with two key issues within regulatory science that need to be dealt with by licensing researchers and regulatory scientists as they think through the best methods for measuring regulatory compliance: 1) Program monitoring paradigms; 2) Relationship of regulatory

compliance/licensing and program quality. The examples drawn are from early childcare and education but the key elements and implications can be applied to any field of study related to regulatory science that involves rules/regulations/standards. For the purposes of this manuscript “rules” will be used to

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

describe or refer to “rules/regulations/standards”.

Program Monitoring Paradigms:

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

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

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

the balance of rules is the goal and solution of any regulatory science paradigm.

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

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

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

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

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

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

compliance, the standards are considered created equally and administered equally.

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

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

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

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

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

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

9) QRIS versus Licensing: examples of a differential/relative regulatory compliance monitoring system would be QRIS – Quality Rating and Improvement Systems. Absolute/full regulatory compliance systems would be state licensing systems. Many programs talk about the punitive aspects of the present human services licensing and monitoring system and its lack of focus on the program quality aspects in local programs. One

should not be surprised by this because in any regulatory compliance system the focus is on "doing no harm" rather than "doing things well". It has been and continues to be the focus of licensing and regulations in the USA. The reason QRIS - Quality Rating and Improvement Systems developed in early care and education was to focus more on "doing things well" rather than "doing no harm". This is not the case in many Canadian Provinces and European countries in which they have incorporated program quality along with specific regulatory requirements.

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

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

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

Regulatory Compliance/Licensing and Quality

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

There are eight basic principles or elements to be presented (they are presented in a binary fashion demonstrating differences): 1) "Do no

harm" versus "Do good". 2) Closed system versus Open system. 3) Rules versus Indicators. 4) Nominal versus Ordinal measurement. 5) Full versus Partial compliance. 6) Ceiling effect versus No Ceiling effect. 7) Gatekeeper versus Enabler. 8) Risk versus Performance.

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

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

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

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

4) Nominal versus Ordinal measurement³. Licensing systems are nominally based measurement systems. Either you are in compliance or out of compliance. Nothing in-between. It is either a yes or no response for each rule. No maybe or partial compliance. With quality systems, they are generally measured on an ordinal level or a Likert scale. They may run from 1 to 3, or 1 to 5, or 1 to 7. There are more chances for variability in the data than in licensing which has 1 or 0 response. This increases the robustness of the data distribution with ordinal measurement.

5) Full or None versus Gradients or Gray Area. Building off of the fourth element, licensing

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

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

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

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

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

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

Reference:

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

Notes:

1. This manuscript should be read along with *A Treatise on Regulatory Compliance* which is referenced above because the two articles build off one another. In the *treatise* description, the specific idiosyncrasies of regulatory compliance data and other key implications of the theory are pointed out that enhance the presentation in this article, such as the extreme nature of skewness that is present in regulatory compliance data, nominal data measurement, the differences between full and substantial regulatory compliance, designing the most cost effective and efficient differential
2. The ultimate goal is the most cost effective and efficient differential monitoring system for negotiating the path of least resistance in complying with a given set of rules which will provide the proper balance of rules. This should be the goal of any regulatory science paradigm. By using the previous *Treatise* article along with this article should provide a blueprint for the regulatory science field in designing a program monitoring system to measure regulatory compliance where an emphasis on differential monitoring should occur in licensing systems and full-scale monitoring should occur in program quality systems. Another approach is to have both regulatory compliance and program quality built as a continuum in the program monitoring system similar to what Head Start is attempting.
3. There are instances in which this dichotomy is not as clear or straightforward where licensing systems do allow partial compliance as a facility has opportunities to correct non-compliances on their way to achieving full compliance with specific rules. The problem is that this is not necessarily a standardized process and it is difficult to determine if it is used often in licensing agencies' monitoring efforts.

A Treatise on the Theory of Regulatory Compliance

Richard J. Fiene^{a,*}

^aEdna Bennett Pierce Prevention Research Center, Pennsylvania State University, 305 Templar Drive, Elizabethtown, Pennsylvania 17022

Abstract

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

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

1. Introduction

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

A factor with regulatory compliance data is that they generally follow a very skewed frequency distribution, which limits analyses to non-parametric statistics. Because of the skewed data distribution, dichotomization of data is warranted, given the lack of variance in the regulatory compliance frequency distribution - the majority of facilities¹ are either in full or substantial regulatory compliance.

An assumption within regulatory compliance is that full regulatory compliance, that is, 100 percent compliance with all rules², is the best (i.e., risk is minimized) possible scenario for the services being delivered and assessed. It is also assumed that all promulgated rules have an equal weight in their relative impact on the desired service delivery model, although this thinking has been changing over time regarding how rules are

reviewed and complied with. This short treatise will examine the past 40 years of research delving into regulatory compliance measurement, and will provide some guidance to regulatory researchers and policy-makers as they move forward with both research and policy development related to rules. The data from these research studies have led to a Theory of Regulatory Compliance that demonstrates that substantial regulatory compliance - and not full regulatory compliance - is a more effective and efficient public policy as it relates to decision making on monitoring and licensing.

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

2. Methods

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

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

¹The term “facilities” is used when referring to programs and/or facilities.

²The term “rules” is used when referring to rules and/or regulations.

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

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

3. Results

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

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

4. Discussion

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

1. Public policies that require full (100 percent) compliance with all rules may not be in the best interest of the clients being served, nor an effective use of limited regulatory resources. Potentially, emphasis on substantial regulatory compliance may be a more effective and efficient public policy related to client outcomes when it comes to their health, safety, and quality of life. Note that substantial compliance is still very high regulatory compliance (99-97 percent compliance with all rules) and produces positive client outcomes. As stated above, regulatory compliance data are extremely skewed and not normally distributed. There is very little variance in the data and the majority of programs are in either full or substantial regulatory compliance.
2. If a jurisdiction focuses on a substantial regulatory compliance public policy it opens up many system enhancements, such as differential or targeted monitoring, risk assessment analysis, and statistical key indicator rules that have been demonstrated to be cost effective and efficient approaches to reviewing program performance. In a full regulatory compliance public policy approach, none of these system enhancements can be employed, with the possible exception of the key indicator approach as delineated in number four below.
3. If a jurisdiction takes the position that all rules are not equal, then a risk assessment or weighting approach becomes an alternative based on the assumption that certain rules place clients at greater risk of death, serious injury, or other types of harm.
4. Even if a jurisdiction does not have a licensing law that allows issuing licenses on the basis of substantial compliance, there is the possibility that key indicators could still be used for abbreviated reviews or inspections, if there is no prohibition in statute or regulation that expressly forbids the use of this approach, since key indicators statistically predict full regulatory compliance. In other words, all rules are statistically predicted to be in regulatory compliance based on the results of the key indicators. Therefore, technically, all rules have been reviewed albeit short of a full review or inspection.
5. Based on previous research, utilizing a risk assessment approach along with a key indicator approach is the most cost effective and efficient differential monitoring system model. The reason is that both predictive rules and those rules that place clients at greatest risk are always assessed when a site visit review or inspection is done. Many more jurisdictions use a risk assessment approach at this point, but there is a loss of predictive regulatory compliance by just using it.
6. Based on previous regulatory compliance history, only those facilities in high regulatory compliance would be eligible for abbreviated key indicator and risk assessment reviews, whereas those with a history of high regulatory non-compliance would continue to receive full regulatory compliance reviews. This gets at the essence of the differential monitoring approach, which is cost neutral. Regulatory resources may then be re-allocated from the abbreviated reviews to more in-depth full regulatory compliance reviews.
7. Based on the use of the key indicator and risk assessment methodologies within a differential monitoring approach, it is possible to identify over multiple jurisdictions if there are generic rules that meet the criteria of risk abatement and prediction. Such an application has occurred in the United States with the creation of early care

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

5. Conclusion

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

6. Declaration of Conflicting Interest

The authors declare no conflicts of interest.

7. Article Information

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

8. References

- [1] Fiene, R. (2013). *Office of Head Start Key Indicator Project Report*. Retrieved from <https://drfiene.files.wordpress.com/2018/04/rheadstart-report.pdf>
- [2] Fiene, R. (2014). *Georgia Child Care Licensing Study: Validating the Core Rule Differential Monitoring System*. Retrieved from <https://drfiene.files.wordpress.com/2018/04/2014-georgia.pdf>
- [3] Fiene, R. (2016). *Early Childhood Program Quality Indicator and Improvement Model (ECPQIM) and Differential Monitoring Logic Model and Algorithm (DMLMA) Readings*. Retrieved from <https://drfiene.files.wordpress.com/2018/03/1ecpqims-book-of-readings.pdf>

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

Richard Fiene, Ph.D.

Affiliated Faculty
Prevention Research Center
The Pennsylvania State University

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

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

Introduction

The purpose of this paper is to compare several countries (N =20) and the United States on the Child Care Aware – formerly NACCRRRA (National Association of Child Care Resource and Referral Agencies) Child Care Benchmarks

that have used extensively in the USA to compare state regulatory and monitoring policy and implementation. The use of these benchmarks has been very useful in comparing states in the USA on an agreed upon series of child care benchmarks that have a great deal of support in the research literature (AAP/APHA, 2012, 2013; NACCRRRA 2007, 2009, 2011). Previous research (OCED, 2006) has focused on early care and education policies in other countries which was a very important

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

first step in making comparisons across countries. This paper will expand upon this comparison in order to begin applying the NACCRRRA benchmarks to other countries and establish a baseline between the USA and other countries related to regulatory review and analysis. This study is important because it provides a common rubric for making comparisons between the USA and other countries that is reliable and valid (NACCRRRA 2007, 2009, 2011)

related to regulatory analysis. As far as the author can determine from his extensive review of the literature, similar studies of this type have not been attempted utilizing a standardized rubric created by a major national child care organization. There have been other studies completed in which comparisons were made of other countries, the OCED (2006) Starting Strong II study and report is an excellent example of this type of

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

$$CI \times PQ \Rightarrow RA + KI \Rightarrow DM$$

Definitions of Key Elements:

CI = Comprehensive Licensing Tool (Health and Safety)(*Caring for Our Children*)

PQ = *ECERS-R, FDCRS-R, CLASS, CDPEs* (Caregiver/Child Interactions/Classroom Environment)

RA = Risk Assessment, (High Risk Rules)(*Stepping Stones*)

KI = **Key Indicators (Predictor Rules)**(*13 Key Indicators of Quality Child Care*)(NACCRRRA Benchmarks)

DM = Differential Monitoring (How often to visit and what to review)

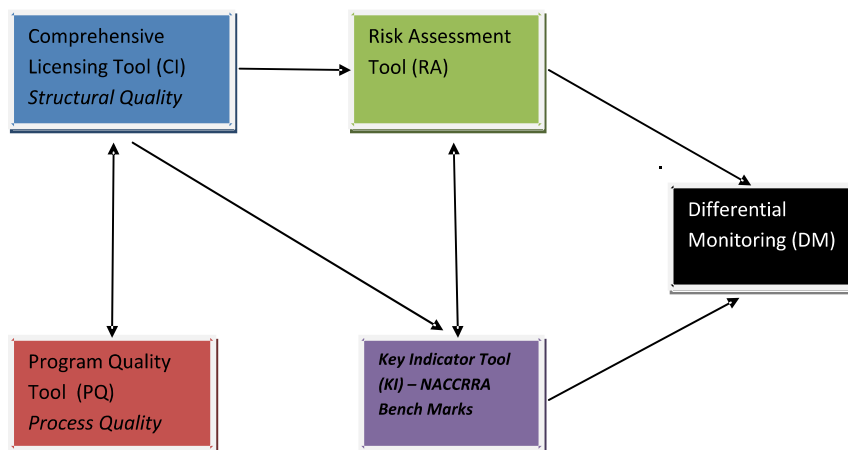


Figure 1.

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

The child care benchmarks¹ utilized in this study are based upon the following key indicators: prevention of child abuse, immunizations, staff child ratio, group size, staff qualifications and training, supervision/discipline, fire drills, medication administration, emergency plan/contact, outdoor playground, inaccessibility of toxic substances, and proper hand washing/ diapering (NACCRRA 2007, 2009, 2011). These benchmarks are more based upon the structural aspects of quality rather than on the process aspects of quality. This is an important distinction between the USA approach and the other countries approaches that becomes important in the explanation of results later in this paper.

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

Method

Data Collection Process

Data collection was done on a 100 point scale which is delineated in Appendix 1 as developed by the Child Care Aware - NACCRRA Research

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

Data Scoring

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

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

Data Collectors

A team of undergraduate and graduate research assistants² at the Pennsylvania State University were the data collectors in which each of them reviewed the child care/early childhood rules/regulations/standards from a specific country and scored the rules/regulations/standards on the Child Care Aware – NACCRRA 100 point raw score protocol and the standardized (0 – 2) scoring approach.

Data Sources

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

Results

The results from this study and analysis were totally unexpected. The results indicated no statistically significant differences between the USA and the

other countries selected (Australia, Belgium, Norway, Finland, Sweden, Ireland, United Kingdom, Italy, France, New Zealand, Mexico, Greece, Canada, Austria, Portugal, Philippines, Turkey, Pakistan, Nigeria, Denmark, and Spain – these countries were selected because of their availability of child care/early care & education rules and regulations as described previously above in Data Sources) when comparing the total scores on the 100 point scale; the USA average for all 50 states scored 58 while the 20 countries average score was 56. However, a very different scenario occurs when looking at the ten individual child care benchmarks using the standardized 0 – 2 scoring protocol. The 20 countries selected in this study scored statistically higher on the following child care benchmarks: Director ($t = 7.100$; $p < .0001$) and Teacher ($t = 7.632$; $p < .0001$) qualifications. The USA scored statistically higher on the following child care benchmarks: Health/Safety ($t = 6.157$; $p < .0001$), Staff Clearances ($t = 3.705$; $p < .01$), and Pre-Service ($t = 4.989$; $p < .001$) /In-Service training ($t = 2.534$; $p < .02$) (See Table 1 & Figure 2).

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

Table 1

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

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

Legend:

Child Care Aware - NACCRRA Benchmarks:

Parent = Parent Involvement (R10)

Health = Health and safety recommendations (R9)

Development = Six developmental domains (R8)

Clearances = Background check (R7)

Inservice = 24 hours of ongoing training (R6)

Preservice = Initial orientation training (R5)

Teacher = Lead teacher has CDA or Associate degree (R4)

Director = Directors have bachelor's degree (R3)

GS = Group size NAEYC Accreditation Standards met (R2)

ACR = Staff child ratios NAEYC Accreditation Standards met (R1)

Scoring:

0.0 = Does not meet Child Care Aware – NACCRRA Benchmarks.

0.5 = Marginally meets Child Care Aware – NACCRRA Benchmarks.

1.0 = Partially meets Child Care Aware – NACCRRA Benchmarks.

1.5 = Substantially meets Child Care Aware – NACCRRA Benchmarks.

2.0 = Fully meets Child Care Aware – NACCRRA Benchmarks.

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

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

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

Discussion

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

comparisons.

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

Legend:

Child Care Aware - NACCRRA Benchmarks:

Parents = Parent Involvement (R10)

Health = Health and safety recommendations (R9)

Devel = Six developmental domains (R8)

Clearance = Background check (R7)

Inservice = 24 hours of ongoing training (R6)

Pre = Initial orientation training (R5)

Teacher = Lead teacher has CDA or Associate degree (R4)

Director = Directors have bachelor's degree (R3)

GS = Group size NAEYC Accreditation Standards met (R2)

ACR = Staff child ratios NAEYC Accreditation Standards met (R1)

Scoring:

0.0 = Does not meet Child Care Aware – NACCRRA Benchmarks.

0.5 = Marginally meets Child Care Aware – NACCRRA Benchmarks.

1.0 = Partially meets Child Care Aware – NACCRRA Benchmarks.

1.5 = Substantially meets Child Care Aware – NACCRRA Benchmarks.

2.0 = Fully meets Child Care Aware – NACCRRA Benchmarks.

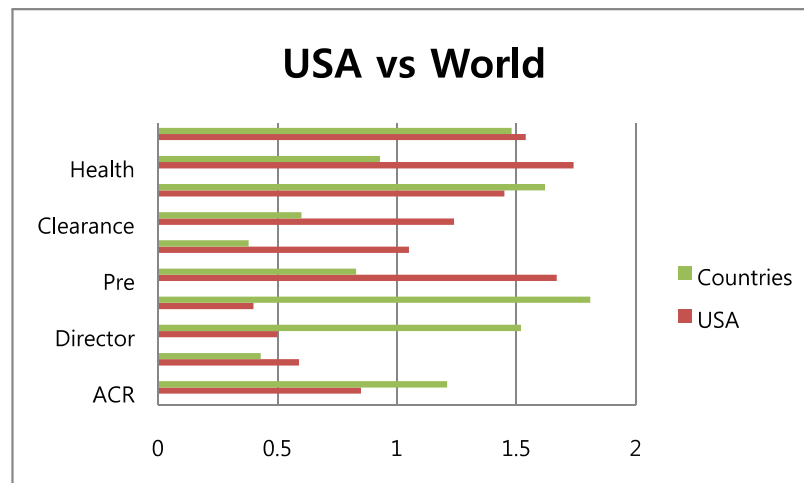


Figure 1. Mean Comparisons between USA and Twenty Countries on Child Care Aware – NACCRRA Benchmarks

rather than the process side of quality.

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

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

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

References

- AAP/APHA (2012). *Caring for our children*, Washington, D.C.: American Public Health Association.
- AAP/APHA (2013). *Stepping Stones*, Washington, D.C.: American Public Health Association.
- Child Care Award - NACCRRA, (2007). *We Can Do Better: NACCRRA's Ranking of State Child Care Center Standards and Oversight*, Washington, D.C.: National Association of Child Care Resource and Referral Agencies.
- Child Care Aware - NACCRRA, (2009). *We Can Do Better: 2009 Update: NACCRRA's Ranking of State Child Care Center Regulations and Oversight*, Washington, D.C.: National Association of Child Care Resource and Referral Agencies.
- Child Care Aware - NACCRRA, (2011). *Leaving Children to Chance: NACCRRA's Ranking of State Standards and Oversight in Child Care Centers*, Washington, D.C.: National Association of Child Care Resource and Referral Agencies

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

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

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

Notes

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

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

Appendix 1

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

Scoring Methods for NACCRRA Ranking of State Child Care Center Regulations (R)										
Question							Scoring method			
Regulation 1. Staff:child ratio requirements comply with NAEYC accreditation standards.							Number of ratios in compliance with NAEYC standards		Score	
							7 ratios		10	
							6 ratios		9	
							5 ratios		8	
							4 ratios		7	
							3 ratios		5	
							2 ratios		3	
							1 ratios		1	
6 mo	9 mo	18 mo	27 mo	3 yr	4 yr	5 yr				
1:4	1:4	1:4	1:4	1:9	1:10	1:10				
R2. Group size requirements are in compliance with NAEYC accreditation standards.							Number of group sizes in compliance with NAEYC standards		Score	
							7 ratios		10	
							6 ratios		9	
							5 ratios		8	
							4 ratios		7	
							3 ratios		5	
							2 ratios		3	
							1 ratios		1	
6 mo	9 mo	18 mo	27 mo	3 yr	4 yr	5 yr				
8	8	8	8	18	20	20				

R3. Center directors are required to have a bachelor's degree of higher in early childhood education or a related field.	<table> <tr> <th>Director education requirement</th><th>Score</th></tr> <tr> <td>Bachelor's degree in any field</td><td>10</td></tr> <tr> <td>College directors certification</td><td>7</td></tr> <tr> <td>Any associate degree</td><td>5</td></tr> <tr> <td>CDA</td><td>5</td></tr> <tr> <td>Clock hours/less than associate degree</td><td>2</td></tr> <tr> <td>High school or less</td><td>0</td></tr> </table>	Director education requirement	Score	Bachelor's degree in any field	10	College directors certification	7	Any associate degree	5	CDA	5	Clock hours/less than associate degree	2	High school or less	0
Director education requirement	Score														
Bachelor's degree in any field	10														
College directors certification	7														
Any associate degree	5														
CDA	5														
Clock hours/less than associate degree	2														
High school or less	0														
R4. Lead teachers are required to have a Child Development Associate (CDA) credential or an associate degree in early childhood education or related field.	<table> <tr> <th>Lead teacher education requirement</th><th>Score</th></tr> <tr> <td>CDA/associate degree or better</td><td>10</td></tr> <tr> <td>State Credential</td><td>5</td></tr> <tr> <td>Clock Hours in ECE</td><td>2</td></tr> <tr> <td>High School/GED</td><td>2</td></tr> <tr> <td>Less than High School</td><td>0</td></tr> </table>	Lead teacher education requirement	Score	CDA/associate degree or better	10	State Credential	5	Clock Hours in ECE	2	High School/GED	2	Less than High School	0		
Lead teacher education requirement	Score														
CDA/associate degree or better	10														
State Credential	5														
Clock Hours in ECE	2														
High School/GED	2														
Less than High School	0														
R5. Lead teachers are required to have initial training, including: • Orientation. • Fire safety. • Other health and safety issues. • At least one staff member certified in first aid must be present when children are in care. • At least one staff member who is certified in CPR must be present when children are in care.	<table> <tr> <th>Number of areas training is required</th><th>Score</th></tr> <tr> <td>Five areas</td><td>10</td></tr> <tr> <td>Four areas</td><td>8</td></tr> <tr> <td>Three areas</td><td>6</td></tr> <tr> <td>Two areas</td><td>4</td></tr> <tr> <td>One area</td><td>2</td></tr> <tr> <td>None</td><td>0</td></tr> </table>	Number of areas training is required	Score	Five areas	10	Four areas	8	Three areas	6	Two areas	4	One area	2	None	0
Number of areas training is required	Score														
Five areas	10														
Four areas	8														
Three areas	6														
Two areas	4														
One area	2														
None	0														
R6. Lead teachers are required to have 24 hours or more of annual training.	<table> <tr> <th>Ongoing training $\geq$</th><th>Score</th></tr> <tr> <td>24 Hours</td><td>10</td></tr> <tr> <td>18 hours</td><td>7</td></tr> <tr> <td>12 hours</td><td>5</td></tr> <tr> <td>6 hours</td><td>2</td></tr> <tr> <td>None</td><td>0</td></tr> </table>	Ongoing training $\geq$	Score	24 Hours	10	18 hours	7	12 hours	5	6 hours	2	None	0		
Ongoing training $\geq$	Score														
24 Hours	10														
18 hours	7														
12 hours	5														
6 hours	2														
None	0														
R7. A comprehensive background check is required for child care providers. • Use of fingerprints to check state records. • Check FBI records. • Check state child abuse registry • Check sex offender registry. • Criminal history check.	<table> <tr> <th>Number of Background checks completed</th><th>Score</th></tr> <tr> <td>Five checks</td><td>10</td></tr> <tr> <td>Four checks</td><td>8</td></tr> <tr> <td>Three checks</td><td>6</td></tr> <tr> <td>Two checks</td><td>4</td></tr> <tr> <td>One check</td><td>2</td></tr> <tr> <td>None</td><td>0</td></tr> </table>	Number of Background checks completed	Score	Five checks	10	Four checks	8	Three checks	6	Two checks	4	One check	2	None	0
Number of Background checks completed	Score														
Five checks	10														
Four checks	8														
Three checks	6														
Two checks	4														
One check	2														
None	0														

R8. Child care centers are required to offer program activities that address all six child development domains • Language/literacy. • Cognitive. • Social. • Emotional. • Physical. • Cultural.	<table><tr><th>Developmental domains addressed</th><th>Score</th></tr><tr><td>6 domains</td><td>10</td></tr><tr><td>5 domains</td><td>9</td></tr><tr><td>4 domains</td><td>7</td></tr><tr><td>3 domains</td><td>5</td></tr><tr><td>2 domains</td><td>3</td></tr><tr><td>1 domain</td><td>1</td></tr><tr><td>None</td><td>0</td></tr></table>	Developmental domains addressed	Score	6 domains	10	5 domains	9	4 domains	7	3 domains	5	2 domains	3	1 domain	1	None	0												
Developmental domains addressed	Score																												
6 domains	10																												
5 domains	9																												
4 domains	7																												
3 domains	5																												
2 domains	3																												
1 domain	1																												
None	0																												
R9. Child care centers are required to follow 10 recommended health and safety practices. • Immunizations. • Guidance/discipline. • Diapering and handwashing. • Fire drills. • Medication administration. • SIDS prevention. • Emergency preparedness. • Playground surfaces. • Hazardous materials. • Incidence reporting.	<table><tr><th>Standards addressed</th><th>Score</th><th>Standards addressed</th><th>Score</th></tr><tr><td>10</td><td>10</td><td>5</td><td>5</td></tr><tr><td>9</td><td>9</td><td>4</td><td>4</td></tr><tr><td>8</td><td>8</td><td>3</td><td>3</td></tr><tr><td>7</td><td>7</td><td>2</td><td>2</td></tr><tr><td>6</td><td>6</td><td>1</td><td>1</td></tr><tr><td colspan="4">Allowing corporal punishment is an automatic zero</td></tr></table>	Standards addressed	Score	Standards addressed	Score	10	10	5	5	9	9	4	4	8	8	3	3	7	7	2	2	6	6	1	1	Allowing corporal punishment is an automatic zero			
Standards addressed	Score	Standards addressed	Score																										
10	10	5	5																										
9	9	4	4																										
8	8	3	3																										
7	7	2	2																										
6	6	1	1																										
Allowing corporal punishment is an automatic zero																													
R10. Child care centers are required to: • Encourage parent involvement. • Require daily or ongoing communication with parents. • Allow parental access any time their children are in care.	<table><tr><th>Number of items required</th><th>Score</th></tr><tr><td>Three items</td><td>10</td></tr><tr><td>Two items</td><td>7</td></tr><tr><td>One item</td><td>3</td></tr><tr><td>None</td><td>0</td></tr></table>	Number of items required	Score	Three items	10	Two items	7	One item	3	None	0																		
Number of items required	Score																												
Three items	10																												
Two items	7																												
One item	3																												
None	0																												

Appendix 2

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



Contents lists available at ScienceDirect

Early Childhood Research Quarterly



Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs

Michel Lahti^{a,*}, James Elicker^b, Gail Zellman^c, Richard Fiene^{a,b,c,1}

^a University of Southern Maine, Muskie School of Public Service, PO Box 9300, Portland, ME 04104-9300, United States

^b Purdue University, United States

^c RAND Corporation, United States

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Child care quality
QRIS
Study design
Methodology

ABSTRACT

In recent years, child care quality rating and improvement systems (QRISs) have become an increasingly popular policy tool to improve quality in early childhood education and care (ECEC) settings and have been adopted in many localities and states. The QRIS proposition is that with higher-quality child care settings, it is more likely that children who attend those high-quality programs will benefit in terms of outcomes like school readiness. However, in order to demonstrate this linkage, QRIS standards and ratings must function as intended, i.e. be valid. This paper presents a framework for validating child care quality improvement standards and processes, along with examples from recent QRIS validation studies in two states. The state examples provide useful data about the strengths and limitations of these validation approaches. We discuss the implications of applying these approaches and provide recommendations to researchers, policy-makers, and program leaders who implement QRIS validation studies.

© 2014 Elsevier Inc. All rights reserved.

Introduction

In recent years, child care quality rating and improvement systems (QRISs) have become an increasingly popular policy tool to improve quality in early childhood education and care (ECEC) settings and have been adopted in many localities and states. The QRIS National Learning Network reports that 40 statewide QRISs have launched or piloted, including the District of Columbia (QRIS National Learning Network, 2014). The immediate goal of a QRIS is to raise the quality of care in early learning settings. Existing research suggests that care in higher-quality settings will improve child functioning, including school readiness (Burchinal et al., 2009; Burger, 2010; Howes et al., 2008), especially for children from lower-income families. QRIS logic models that guide these large-scale interventions focus on improving various dimensions of ECEC quality, with the ultimate goal of improving system outcomes, namely; child care program quality, training and technical assistance for child care providers, information and support for families,

and, therefore, improvements to children's cognitive, language, social, emotional, and physical development.

The perceived need for QRIS has grown out of documented gaps in quality in existing ECEC programs, especially those serving children from lower-income families (Fuller, Loeb, Kagan, & Carrol, 2004; NICHD ECCRN, 2000) and the inability of the current ECEC system to promote uniformly high quality (Cochran, 2007). QRISs produce program-level quality ratings based on multi-component assessments designed to make ECEC quality transparent and easily understood to parents and other stakeholders. Most also include feedback, technical assistance, and incentives to both motivate and support providers' efforts toward quality improvement (Tout et al., 2010). To make program quality transparent, QRISs typically rely on a multi-tiered rating system with one to five levels of program quality. Therefore, it is important that these ratings show evidence of validity, so that higher-quality programs are rated higher, and lower-quality programs are rated lower.

Recent research has documented the importance of both specificity and thresholds when testing hypotheses about child care quality impacts on children's developmental outcomes (Burchinal, Peisner-Feinberg, Bryant, & Clifford, 2000; Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Howes, Whitebook, & Phillips, 1992; NICHD ECCRN, 2000, 2002). However, common global measures of classroom quality such as the Early Childhood Environment

* Corresponding author. Tel.: +520-326-5154; fax: +520-326-5155.

E-mail addresses: mlahti@usm.maine.edu, Michel.lahti@lecroymilligan.com (M. Lahti).

¹ Professor Emeritus at Pennsylvania State University.

Rating Scale-Revised (ECERS-R; Harms, Clifford, & Cryer, 2005) are not always significantly associated with specific child outcomes (Burchinal, Kainz, & Cai, 2011). This may be because these global quality scales do not focus enough on the particular child care quality processes most likely to bring about improved child outcomes (specificity) or they do not provide guidance for the level of quality required to produce improved child outcomes (thresholds). As states implement QRISs, they are using observational measures such as the ECERS-R, and they may also combine other quality measures such as the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008) or locally specified quality indicators. Because QRIS quality standards are often complex, including many components and measures at several quality levels, and because they vary from state to state, it is especially important for states to carefully validate their quality rating systems and match measures specifically to the stated outcome goals of the QRIS. For example, if a particular QRIS places more emphasis on the health aspects of children's development, then the ECERS-R and CLASS would not be appropriate tools; but a tool measuring child care health indicators, such as the National Health and Safety Tool being developed by the California Child Care Health Program (Alkon, 2013) would be more appropriate.

Validity data can also enable researchers to test conclusions about whether the quality indicators embedded in QRIS standards lead to adequate quality assessment and whether the methods used to assign quality ratings are working as intended (Cizek, 2007). This paper defines operationally the concept of QRIS validity, presents four general approaches to assessing validity in the context of large-scale QRISs, and critically examines the efforts of two states, Maine and Indiana, to assess the validity of recently implemented QRISs using these approaches.

Validation of a QRIS is a developmental and multi-step process that assesses the degree to which design decisions about program quality standards and measurement strategies are resulting in accurate and meaningful quality ratings. Validation of a QRIS provides designers, administrators, and stakeholders with crucial data about how well the system is functioning. A carefully designed plan for ongoing QRIS validation creates confidence in the system and a climate that supports continuous quality improvement at both the child care provider and system levels (Zellman & Fiene, 2012).

To date, QRIS validation research efforts have been limited, for a number of reasons. First, validation is complex and involves a range of activities, which should include validating standards, measures, and rating protocols. Second, there has been little information available in the field that clarifies the importance and purpose of QRIS validation or identifies recommended strategies. Third, child care quality advocates and policy makers have been extremely busy designing and implementing these statewide systems, often with limited resources. Given these constraints, validation may seem like an abstract luxury that can wait until later. Further, in states with more mature QRISs, there may be some reluctance among stakeholders to assess the validity of an established and accepted quality improvement system. In newer state systems, policymakers may question the need for validation, given arguments recently offered in support of establishing a QRIS system (Zellman & Fiene, 2012; Zellman, Brandon, Boller, & Kreader, 2011). Yet early and ongoing validation research is essential to the long term success of any system.

One challenge is that QRIS validation cannot be determined by a single study. Instead, validation should be viewed as an iterative process with several equally important goals: refining the QRIS quality standards and ratings, improving system functioning, and increasing the credibility and value of rating outcomes and the QRIS system as a whole. A carefully designed validation plan can promote the accumulation of evidence over time that will provide a sound theoretical and empirical basis for the QRIS (AERA, APA, & NCME,

1999; Kane, 2001; Zellman & Fiene, 2012). Ongoing validation activities, carried out in tandem with QRIS monitoring activities (those that examine ongoing implementation processes) and evaluation activities (those that examine specific outcomes) can help a QRIS improve throughout its development, implementation, and maturation (Lugo-Gil et al., 2011; Zellman et al., 2011).

QRIS validation research may produce three important benefits. First, validation evidence can promote increased support for the system among parents, ECEC providers, and other key stakeholders. Ratings that mirror the experiences of parents and providers can build trust and increase the overall credibility of the system. Second, a system that is measuring quality accurately and specifically should better be able to target limited quality improvement resources to programs and program elements most in need of improvement. This should result in more targeted and effective supports for programs striving to offer higher-quality services. Third, validation evidence can be used to improve the efficiency of the rating process. If a QRIS is expending resources to measure a component of quality that is not making a unique contribution to a summary quality rating, is not measuring quality accurately, or is not contributing to desired program outcomes, that component can be removed or revised. For example, measures that vary little across providers, whose quality varies substantially in other ways, make little or no contribution to overall quality ratings (Zellman & Fiene, 2012).

Four approaches to validation

A comprehensive QRIS validation plan includes multiple studies that rely on different sources of information and ask different but related questions. We suggest QRIS validation research be organized around four complementary approaches: key quality concepts; quality measurement; ratings outputs; and links to child outcomes (Zellman & Fiene, 2012). Summaries of these approaches are provided in Table 1, which includes the purpose of each validation approach, the types of research that can be undertaken, the questions that are asked, and some limitations of each approach. The four approaches are also elaborated later in the paper, as we summarize results of validation research in Indiana and Maine.

In reviewing the table, and throughout this paper, we use three key QRIS terms: component, standard, and indicator. The term 'quality component' refers to broad quality categories used in QRIS (such as staff qualifications, family engagement, or learning environment). A 'quality standard' is defined as a specific feature of quality, such as specialized training in the use of developmentally appropriate curriculum or developmental assessment training within the staff qualifications component. A set of quality standards comprise each quality component. 'Quality indicators' are the specific metrics used for each quality standard. A given quality standard may have one or more quality indicators. An indicator related to the curriculum/assessment staff training standard may be, for example, "At least 50% of teaching staff have completed the two-course statewide training session on developmentally appropriate curriculum."

QRIS validation in Indiana and Maine

This section will describe efforts at QRIS validation in two states in order to explore current validation efforts using these four approaches and to identify the successes and challenges experienced in these early QRIS validation studies. In Indiana and Maine, the QRIS designs are similar, but some aspects of the states' child care contexts, specific QRIS quality components, standards, and rating processes employed are somewhat different. Both states launched their QRIS statewide in 2008, and both systems have four quality tiers, referred to as "levels" in Indiana and "steps" in Maine, organized into a "building block" framework, meaning that child

Table 1
Four related approaches to validating a QRIS.

Approach	Activities and purpose	Typical questions	Issue and limitations
1. <i>Examine the validity of key underlying concepts.</i>	Assess whether basic QRIS quality components and standards are the “right” ones to include by examining levels of empirical and expert support.	*Do the quality components capture the key elements of quality? *Is there sufficient empirical and expert support for including each standard?	*Process subject to interpretation and to political pressure *Limited empirical evidence available; few established links to outcomes of interest.
2. <i>Examine the measurement strategy and psychometric properties of measures used to assess quality.</i>	*Examine properties of key quality measures, e.g., inter-rater reliability on observational measures, scoring of documentation, and inter-item correlations, to determine if measures are psychometrically sound. *Examine relationships among quality measures to assess whether they function as expected.	*What is the reliability and accuracy of indicators collected using different methods? *Do quality measures perform as expected? (e.g., do expected subscales emerge?) *Do measures of similar standards relate more closely to each other than to other measures? *Do measures relate to each other in theoretically consistent ways?	*This validation activity is especially important given that some quality component measures were likely developed in low-stakes settings and have not been examined in the high-stakes context of QRIS.
3. <i>Assess the outputs of the rating process</i>	*Examine variation and patterns of program-level ratings within and across program types, to assess if QRIS distinguishes levels of quality. *Examine relationship of program-level ratings to other validated quality indicators to determine if ratings are assessing quality in expected ways. *Examine alternate cut points and combining rules to determine how well the ratings distinguish different levels of quality.	*Do rating distributions vary by program type, e.g., center-based programs vs. home-based programs? *Do programs with different program-level ratings differ in meaningful ways on alternative quality measures? *Do levels cut scores and combining rules produce expected rating distributions and meaningful distinctions among programs?	*Measurement error is an important issue that should be examined. *These validation activities depend on a reasonable level of confidence about the quality components, standards and indicators as well as the process used to designate ratings. *Comparing QRIS measures to other measures is frequently constrained by the absence of validated alternative measures of the same constructs.
4. <i>Examine how ratings are associated with children's outcomes.</i>	Examine the relationship between program-level ratings and selected child outcomes to determine whether higher program ratings are associated with better child outcomes.	Do children who are attending higher-rated programs have greater levels or gains in skills than children who attend lower-quality programs?	*Appropriate demographic and program level control variables must be included in analyses to account for selection factors. *Studies may be done on selected child and program samples to save resources. *Given staff turnover and movement, children may not spend substantial time with quality-rated caregivers (dosage.) *Measurement limitations noted for Approach 3 apply here as well. *Findings from non-experimental studies do not permit attribution of causality about QRIS participation, but preliminary inferences can be made about how quality influences children's outcomes.

care providers must enter at the lowest level and meet all quality standards and indicators at each level in order to advance to the next higher level. The focus on these two states in this paper is to help illustrate the application of these four approaches to operationalizing validation in a QRIS. While the QRIS evaluations in Maine and Indiana have resulted in other kinds of information disseminated for policy makers in these states and publications for other audiences, this paper is unique in that it is only intended to focus on these four concepts of validation.

Both states partnered with university-based researchers to conduct validation research, after piloting aspects of their QRIS design. However, there are also key differences between these two states. For example, the Indiana QRIS standards were developed based on a local community-based model that was then modified by a state stakeholder committee for statewide expansion. The Maine quality standards were developed to align with program-type-specific national accreditation standards. The Maine

standards were also vetted through review and comment by many stakeholders and technical assistance was provided by University researchers based on reviews of the scientific literature. Maine QRIS ratings are generated by provider self-report, then verified by state agency staff, while Indiana employs independent raters who directly assess the standards by visiting child care settings. Provider voluntary participation rates are higher among state-licensed providers in Indiana. However, Indiana also has significant numbers of license-exempt child care providers, whereas license exemption is not a prominent feature of the Maine child care system. The key features of each state QRIS are summarized in Table 2. These two states provide useful examples, because while the state child care contexts are different, they each used strategies contained in the four validation approaches discussed above and outlined in Table 1. The successes and limitations of these states' approaches will inform future validation research on QRIS.

Table 2
Key features of Indiana and Maine QRISs.

QRIS feature	Indiana	Maine
System name	Paths to QUALITY (2008)	Quality for ME (2008)
Eligible child care types (participation rate)	Licensed centers (89%) Licensed homes (62%) Unlicensed registered ministry centers (12%)	Licensed centers (68%) Licensed homes (44%)
Participation rules	Voluntary for all providers (Not required for CCDF participation)	Required for programs participating in federal CCDF/Voluntary for all others
QRIS structure	Building Block; 4 quality levels	Building Block; 4 quality levels
QRIS standards (examples)	Level 1 – Licensed, or completes voluntary certification program Level 2 – Learning environment and materials requirements; daily literacy activities; 25% of staff have CDA or equivalent; 15 hrs. in-service training/yr.; etc. Level 3 – Written curriculum focused on whole child; provision for special needs; 50% of staff have CDA or equivalent; 20 hrs. in-service training/yr.; etc. Level 4 – National accreditation; Provide mentoring to other QRIS providers (see www.in.gov/fssa/carefinder/2554.htm#).	Step 1 – Meets all regulatory standards, in operation for more than one year, and all staff registered in Maine Roads to Quality Registry (MRTQ). Step 2 – Learning Environment/Developmentally Appropriate Practice requirements; program improvement plan in place; 50% of staff at level 5 on MRTQ career lattice; etc. Step 3 – Documented use of Early Childhood Learning Guidelines and/or Infant-Toddler Learning Guidelines; Evidence collected at least three times per year on child's development; etc. Step 4 – National accreditation; written parent involvement plan; etc. (See: https://www.maine.gov/dhhs/ocfs/ec/occhs/qualityforme.htm).
QRIS standards development process	Aligned with national accreditation standards, expansion of community pilot program, modified by state stakeholder committee.	Aligned with national accreditation standards; based on available empirical evidence; and from results of stakeholder reviews and comment.
QRIS rating procedure	Independent ratings contractor, annual site visits, using Paths to QUALITY standards checklist	(1) Provider self-assessment, online system; (2) enrollment system uses linked files from state licensing and registry; and (3) verified by state agency staff.

Method

Indiana

The Indiana QRIS is called “Paths to QUALITY™.” The validation research reported here includes a preliminary literature review and an empirical field study including a stratified random sample of 276 child care providers who had voluntarily entered the QRIS during 2008–2009, including 135 classrooms in 95 licensed child care centers, 169 licensed family child care homes, and 14 classrooms in 12 unlicensed registered child care ministry centers. Independent, on-site assessments were completed by university researchers approximately one year after QRIS entry and included: observational global quality assessments of the child care environment using the Environmental Rating Scales (ERS: ITERS-R, ECERS-R, FCCERS-R; Harms et al., 2005); observations of adult–child interaction quality (Caregiver Interaction Scale, CIS; Arnett, 1989); surveys and interviews with child care providers; and interviews with parents whose children had been placed with QRIS providers. Observers were trained to reliability level of 80% exact agreement ($Kappa = .70$) or higher, and maintained reliability during the study. Child development assessments were completed using standardized research-validated measures, with two randomly selected children from each participating child care center classroom or family child care home. For children under three years, measures included the Mullen Scales of Early Learning (Mullen, 1995) for cognitive and language development and the Brief Infant Toddler Social Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2002) for social-emotional development. For children three to five years, the measures included the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 1997) and Woodcock–Johnson Applied Problems and Letter–Word Identification subtests (Berry, Bridges, & Zaslow, 2004) for language and cognitive development and the Social Competence and Behavior Evaluation (SCBE; LaFreniere & Dumas, 1997) for social-emotional development. (For a detailed description of the Indiana evaluation methodology, see Elicker et al., 2013; Elicker, Langill, Ruprecht, Lewsader, & Anderson, 2011.)

Maine

The Maine QRIS is called “Quality for ME.” The Maine validation research reported in this paper is based on a literature review of quality variables, focus group interviews with providers and parents, and a field study including a stratified random sample of 255 providers who enrolled in the QRIS in 2008 through 2011, including: 153 classrooms in 105 licensed child care centers; 113 licensed family child care homes; and 41 classrooms in 37 Head Start sites. Assessments were completed as soon as possible after a program enrolled into the QRIS, however, this varied based on the length of time required for the state agency to verify enrollment information and schedule on-site observations. Like Indiana, Maine researchers used the ERS global quality assessment scales and conducted surveys with providers and parents. Unlike Indiana, Maine did not collect any child-level outcome data. Assessors were trained to reliability annually by authors of the ERS scale and maintained a 85% inter-rater reliability during the study. (For a detailed description of the evaluation methodology, see Lahti et al., 2011.)

Results

Results of the QRIS validation research in Indiana and Maine are presented in relation to the four approaches to validation recommended by Zellman and Fiene (2012; refer to Table 1).

Approach 1: examine the validity of key underlying concepts

As noted above, the quality components included in a QRIS (e.g., staff qualifications, learning environment, family engagement) essentially define how child care quality will be viewed in each state. Conceptual validation provides justification and support for these chosen elements. This first validation approach asks whether the quality components, standards, and indicators included in a QRIS are the “right” ones; that is, if together they define quality of care. Many state QRISs have adopted similar, though not identical, concepts and program quality standards (Smith, Robbins, Stagman, & Kreader, 2012).

One approach that can help to validate the underlying concepts of quality in a QRIS involves assessing the degree to which the quality components used in the QRIS ratings include standards and indicators that are based on empirical evidence that links them to desired program, family and child outcomes. A literature review weighs the existing research evidence and on that basis provides a judgment about whether a particular quality component should be included or excluded from the QRIS. Like many validation activities, such reviews ideally would be updated from time to time to determine if revisions to the QRIS are advisable in light of new research findings. As noted in Table 1, this approach may be limited by available data. Further, available data may be subject to more than one interpretation. Politics can also play a role; supporters of particular elements, e.g., nutrition, accreditation, may want to ensure that such measures are included, regardless of the strength of the research evidence. This literature review approach of conceptual validation was a key method used in developing both Indiana's and Maine's QRIS quality standards.

Indiana: examining the validity of underlying concepts

Standards and indicators for each QRIS level in Indiana were drafted by a state committee of child care providers and stakeholders. The standards were based on an existing community-level Paths to QUALITY model, but also made accommodations for statewide use and integration into the existing state child care licensing and training/technical assistance systems. The highest level quality goal for QRIS in Indiana is national accreditation, so proposed quality standards and criteria at each QRIS level were constructed to help child care providers work toward accreditation in steps.

The Purdue University research team conducted a review of previous evaluations of the Indiana QRIS community-level pilot programs and an in-depth analysis of the proposed QRIS quality standards based on the published child development and child care literature. This literature-based analysis is summarized here. (For a full report, see Elicker, Langill, Ruprecht, & Kwon, 2007; Elicker et al., 2013.)

First, the evaluators looked at each proposed QRIS quality standard and indicator for each type of child care. Ten broad quality components were identified that encompassed all of the proposed quality indicators: regulation; teacher education/training; structural/environmental quality; process quality/interactions; assessment; provisions for children with special needs; program policies; director/owner professional development; parent–teacher communication; and national accreditation. These ten components were then used as key terms to guide an extensive search of the research literature to collect and weigh the available evidence that each component was: (1) generally considered a valid aspect of quality; and (2) empirically associated with children's well-being or positive developmental outcomes. Based on the amount and quality of evidence, each quality component was rated as follows: (1) some or limited evidence (one or two well-designed studies); (2) moderate evidence (3–5 well-designed studies); or (3) substantial evidence (more than five well-designed studies). The results of this analysis were reported to the state QRIS planning committee, including a conclusion that most of the proposed quality indicators had “substantial evidence” for their validity.

Maine: examining the validity of underlying concepts

Researchers at the University of Southern Maine worked with state agency leaders and other key stakeholders through a process that involved the use of Concept Mapping (The Concept System®, 2012). This process allows for the development of a conceptual framework that can guide planning, and in this case led to the selection of the underlying quality concepts and standards for Maine's QRIS. Similar to what was done in Indiana, University of

Southern Maine research staff identified key quality concepts from the literature and national accreditation standards. In addition, concepts emerged from results of eight focus groups with parents and ECEC professionals across the state, including participants from various types of settings, e.g., family child care homes, after school programs, centers, etc. Statements of program quality were developed; these statements were the focus of a mapping process which involved more than twenty-four experts reviewing and rating the statements. The Concept Maps that result from this process allowed participants to visually identify which concepts of program quality were most favored by specific key stakeholder groups. In addition, the mapping software illustrated how closely related the concepts were to each other, based on reviews from the select experts. From this process, a set of components and standards was developed. The final step in selecting program quality standards involved a formal review and comment process that the state agency implemented in various locations across the state (Maine DHHS, 2008).

Approach 2: examine measurement strategies and psychometric properties of quality measures

A second type of validation focuses on the attributes of the individual quality measures used in the QRIS and the way these measures are combined to produce a summary rating of program quality. This approach addresses how well measures are working in the context of the QRIS. These efforts attempt to answer questions such as, “Is there evidence that a given indicator measures what it purports to measure?” “If the QRIS claims to have a specific number of dimensions, do we find those dimensions in the output data?” “Is there sufficient variance in scores on this quality indicator to justify its inclusion in the QRIS?” Addressing these issues involves an examination of the distribution of participating provider quality scores and the internal consistency of multi-item measures.

The research literature provides limited guidance concerning the most appropriate ways to combine measures of quality indicators into summary ratings (Lugo-Gil et al., 2011; Tout, Zaslow, Halle, & Forry, 2009; Zellman, Perlman, Le, & Setodji, 2008). Yet this process is crucial to producing meaningful overall program quality ratings, the key output of the rating assessment process. At minimum, it is important to consider whether certain elements *should* be treated as more important, and if so, how this can be assured in the process of combining them. If this issue is not addressed, unexamined weighting may occur anyway. For example, if measures of individual quality elements are combined without any weighting, then those measures that are longer (e.g., include more items) will count for more in a final rating.

At the time both the Maine and Indiana QRISs were being designed, in the mid-2000s, the predominant global quality measures in use in both states were the Environmental Rating Scales (ERS) (Harms et al., 2005; Harms, Cryer, & Clifford, 2006; Harms, Cryer, & Clifford, 2007). ERS use was predominant in the accreditation quality improvement efforts in both child care centers and child care homes. So there was some familiarity with the measures on the part of providers. This was an important political consideration in terms of developing and promoting the design of the QRIS. In addition, in reviewing emerging QRIS work from other states, it appeared that the ERSs were the predominant global classroom quality measure in use at that time. While ERS was influential in the design of the QRIS quality standards in both Maine and Indiana, it is important to note that the ERS are not used to determine the step or level quality ratings. Many other quality indicators are included in the QRIS standards of both states, including staff qualifications, annual staff training hours, and other indicators that help providers make progress toward the ultimate quality goal of national accreditation.

Choosing the points at which individual measures (in block design QRISs) and summary ratings are assigned to rating levels is another exercise that has received limited attention. Cut scores can be assessed in a number of ways. One relatively simple one is to use existing data to conduct a “virtual pilot” (Zellman & Karoly, 2012a) in which existing data are used and cut scores are altered and the effects are examined in terms of distributions of summary ratings across programs. A downward limit on cut scores is the need for some variation within each quality component; without it, a component provides no useful information in overall ratings. Designers may compare program distributions using different cut scores, although it is not always clear what an appropriate rating levels distribution should be. However, it is reasonable to assume that an appropriate distribution in the early phase of a QRIS would be one in which there are programs placed at all levels, with decreasing numbers of programs at each succeeding higher level.

Another validation activity might involve an assessment of the relationship of a given indicator to other indicators of quality included in the QRIS. In studies that examine measures to be included together in a QRIS, it is important to look at the degree of correlation found among these measures: ideally, measures will be moderately correlated so that each measure both contributes to an overall assessment of quality yet also provides some non-redundant program quality information (Zellman et al., 2008). Correlation patterns should make sense. For example, two measures of interaction quality should be more closely related to each other than to a measure of adult–child ratios. If such studies reveal for example that the correlation between ratios and interaction processes is very high ($r = .90+$) this result might argue for eliminating one or the other indicator from the QRIS, as they may not be providing unique information (although some QRISs include certain quality elements to ensure that they are paid attention to for other policy related reasons, even if their psychometric properties are not ideal). To date, the Maine and Indiana validation research has not included a comparison of measures internal to the QRIS rating systems, but this is recommended in future research as the systems mature and stabilize.

Measurement error presents another potential challenge in assessing QRIS validity. Most QRISs assume that observational measures are relatively stable over time absent quality improvement efforts. This assumption is consistent with empirical evidence for at least one widely used instrument, the ERS (Clifford, 2005).

A related measurement issue concerns inter-rater reliability. In the twenty systems reviewed by Tout et al. (2010), nearly all QRISs require 80–85% agreement with a master coder (either exact agreement or agreement within one scale point) on ERS; this degree of reliability does not eliminate errors in ERS measurement (Bryant, 2010; Bryant, Burchinal, & Zaslow, 2011). For instance, two raters could be 100% reliable under a standard of 85% agreement within one scale point, but one might give a classroom a score of 3.5 and the other a score of 4.5, a difference that is large enough to affect an overall program rating (Karoly, Zellman, & Perlman, 2013). Based on the range and degree of variability in ERS quality scores at each rated level in both Maine and Indiana QRISs (see Tables 3 and 4), we recommended that program managers strive to increase the reliability of the rating process by clearly defining quality indicators and rating procedures, and conducting regular reliability checks.

Approach 3: assess the outputs of the rating process

A third validation approach focuses on assessing the *outputs* of the rating system: the scores and levels assigned to providers who undergo a rating, and the distributions of those scores within and across different types of providers. Studies conducted under this approach examine the degree to which the quality levels in the QRIS are meaningfully distinct from each other. The results of these

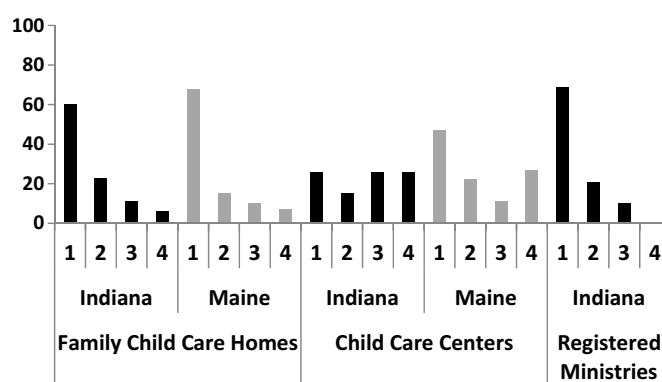


Fig. 1. Percent enrolled providers rated at four quality levels, by state and type of child care.

studies may provide data that suggest that measures, cut scores, or rules for combining measures need to be changed in order to distinguish the rated quality levels effectively. Because these studies can result in proposals for significant changes to the standards for QRIS levels, it is helpful for these studies to occur prior to studies that examine associations between quality levels and children's development.

Output studies may focus on individual indicator scores, such as how providers score on an environmental rating, as well as on the overall quality level that is the final output of the rating process. These studies may also utilize a measure of quality not included in the QRIS rating process to make an evaluation of concurrent validity, by examining whether assessments on both measures co-vary in predictable ways. The following section provides examples of the two states' examinations of the distribution of quality ratings and rating-level advancement patterns for each program type enrolled in the state QRIS.

Examining initial QRIS rating distributions and cut points

While evaluators in Indiana and Maine did not conduct a detailed examination of the weighting or internal consistency of specific quality indicators, they did analyses to reveal the distribution of quality levels. After three years of system implementation, both Indiana and Maine QRIS child care providers were predominately rated at Level 1 or Level 2 (see Fig. 1). It is important to note that in Indiana, all providers enter the system at Level 1, and in Maine, providers can enter the system at any level based on their program rating, and then may advance at will from that level. A recent in-depth study of five state quality rating and improvement systems that were fully implemented found a similar pattern, with four of the five states reporting 40–76% of all programs enrolled in the lower tiers of the system (Mathematica Policy Research, 2011).

In Indiana, licensed child care centers were evenly distributed across the four QRIS levels approximately two years after the program inception. However licensed family child care homes were most frequently found at Level 1, with steeply declining numbers at the other three levels. This higher proportion of Indiana licensed centers rated at Level 3 or Level 4 may have been due to a greater historical emphasis in child care centers than in homes on regulation and attaining national accreditation, greater organizational capacity to complete the requirements of advancement in QRIS, or possibly that QRIS standards more closely reflect center quality than family child care home quality. Unlicensed registered child care ministries, a unique type of child care center in Indiana that is not licensed due to religious affiliation, participated at a much lower rate, and none had yet attained Level 4, reflecting significant challenges facing these unlicensed centers in meeting the Level 1 standards needed to enter the QRIS. These data patterns in Indiana

Table 3Indiana QRIS: mean global quality ERS scores^a as a function of program type and rated quality level.

	Level one (n = 84)	Level two (n = 90)	Level three (n = 74)	Level four (n = 66)
All providers (N = 314)	3.2 (.87)	3.7 (.76)	3.8 (.73)	4.3 (.80)
Family child care homes (n = 167)	2.9 (.64)	3.4 (.75)	3.6 (.67)	4.0 (.89)
Licensed child care centers (n = 133)	4.0 (.77)	4.0 (.68)	4.3 (.66)	4.5 (.67)
Unlicensed registered child care ministries (n = 14)	3.2 (.95)	4.1 (.45)	4.0 (.18)	NA

^a Possible range = 1–7.**Table 4**Maine QRIS: mean global quality ERS scores^a as a function of program type and rated quality level.

	Step one (n = 82)	Step two (n = 99)	Step three (n = 79)	Step four (n = 82)
All providers (N = 342)	3.7 (.77)	3.9 (.84)	4.0 (.80)	4.3 (.79)
Family child care homes (n = 129)	3.3 (.67)	3.5 (.80)	3.8 (.91)	4.2 (.83)
Licensed child care centers (n = 165)	3.9 (.72)	4.1 (.80)	4.2 (.68)	4.4 (.84)
Head start centers (n = 48)	NA	NA	4.1 (.75)	4.5 (.71)

^a Possible range = 1–7.

supported the validity of the QRIS rating system in that they showed variation in quality ratings across participating providers, they reflected the increasing effort necessary to meet quality standards at higher levels, and they were interpretable within the state's child care context.

Another gauge of overall quality rating system utility is the amount of program advancement to higher rated levels. It is reasonable to expect, if the QRIS is viable, that at least some providers will advance in quality level. In the Indiana evaluation, 19% of the licensed centers, 24% of the licensed homes, and 27% of the unlicensed ministry centers had advanced at least one QRIS quality level in a 6-month period between assessments, during which mentoring was provided by local training providers. This advancement pattern, if maintained over time, suggests that even though attaining the highest levels may be challenging, quality improvement is feasible.

For Maine, as Fig. 1 illustrates, center-based programs and family child care type programs are most frequently found at Step or Level One. A disproportionately small number of family child care programs have attained Step Four, the highest quality level, and a disproportionately large number of child care centers and Head Start programs are enrolled at Step Four. This pattern of fewer family child care homes enrolled at higher Step levels has existed throughout QRIS implementation in Maine. Maine family child care home providers argued that some of the program standards were not “a good fit,” despite designers' beliefs that standards were well-matched to setting type. The large number of center-based and Head Start programs at the higher Step levels was expected, given that QRIS quality standards closely align with accreditation standards, and center-based programs are more likely to be nationally accredited than family child care homes.

An assumption of the designers of the QRIS in Maine was that programs engaged with QRIS will improve their tier levels consistently over time (Lahti et al., 2011). Approximately 80% of all programs (n = 1118) in the QRIS observed during the study period 2008 through 2011 did not experience a move up from one Step Level to the next. Results indicated that 95 of the 103 events or changes in Step Level from level one to two occurred during the first 23 months of enrollment. Moving from a Step One to Two, center-based care programs had a hazard probability of just .02 while family child care homes stayed virtually flat during this early period of enrollment in the QRIS. For movement from Step Two to Three, neither program type (p = 290) nor regional location (p = 195) appear to be significant in explaining Step level movement. For movement from Step Three to Four, the highest tiers in Maine's QRIS, the analysis indicated that only type of program is a significant covariate explaining advancement. Family child care homes

appeared to have a significantly lower probability of advancing a Step Level at this highest quality tier, compared with center-based and Head Start programs. These types of analyses of program movement in the system are relevant to the validation of a QRIS as they illustrate whether or not the way the system as designed is meeting its goals of supporting program advancement, leading to statewide improvement of program quality.

Studies may also be conducted to examine the degree to which given measures relate to other measures that purportedly assess the same concept. Here, strong correlation is desired, as they suggest that measures are measuring the concepts that they purport to measure in ways that are consistent with other measures of the same concepts.

Indiana: assessing the output of the rating process

The Indiana evaluation research included one validation test of state committee-generated quality standards, indicators, and levels by comparing the outputs of the QRIS rating system with independently gathered assessments of quality using validated quality measures, the environmental rating scales (ERS, Harms et al., 2005, 2006, 2007) and the Caregiver Interaction Scale (CIS; Arnett, 1989). The results, originally published by Elicker et al. (2011) and shown in Table 3, indicate that ERS scores co-varied as expected with QRIS level ratings, with a significant mean difference in global scores of 1.1 scale points between Level 1 and Level 4. Table 5 shows that caregiver interaction as observed using the CIS was less related to the rated QRIS quality levels. The overall correlation between the 4-level QRIS ratings and global ERS quality scores was moderate ($r = .42$, $p < .01$). The correlation between CIS adult-child positive interaction scores and QRIS level was more modest, but positive and significant ($r = .24$, $p < .01$).

Taken together, and looking across all types of providers, these results suggest that the QRIS ratings distinguish levels of quality in somewhat similar ways as two time-tested, validated measures of child care quality. However, mean quality levels at Level 4 were mostly found to be below the “good” rating threshold, suggesting the need to strengthen standards and/or rating procedures at the highest QRIS levels. In addition, finer analysis of the data suggested specific recommendations about quality standards and rating procedures that might be improved for each type of child care. Summaries of individual ERS item means for Level 3- and 4-rated providers led to the identification of a number of ERS items with scores below 4. Program planners are currently improving standards and QRIS rating procedures in light of these findings (Elicker et al., 2013).

In Indiana, patterns of association between QRIS ratings and ERS ratings were not the same for all types of child care. While the

Table 5Indiana QRIS: mean total positive interaction (CIS) scores^a as a function of program type and rated quality level.

	Level one (n = 84)	Level two (n = 87)	Level three (n = 74)	Level four (n = 65)
All providers (N = 310)	3.0 (.44)	3.1 (.46)	3.2 (.37)	3.2 (.52)
Family child care homes (n = 164)	2.9 (.37)	3.1 (.43)	3.2 (.36)	3.3 (.48)
Licensed child care centers (n = 132)	3.2 (.44)	3.1 (.50)	3.2 (.41)	3.2 (.54)
Unlicensed registered child care ministries (n = 14)	2.8 (.58)	3.2 (.48)	3.5 (.11)	NA

^a Possible range = 1–4.

global ratings were significantly correlated in both licensed centers and licensed family child care homes, the strength of association was stronger for homes, meaning QRIS level ratings in homes more clearly distinguished levels of ERS-related quality, meaning at each QRIS-rated level, the ERS quality differences were generally greater than they were for centers. Second, the overall ERS quality levels for center-based preschool classrooms (using ECECRS-R; $M = 4.6$ at Level 4) were somewhat higher than for center-based infant classrooms (ITERS-R; $M = 4.4$ at Level 4) and family child care homes (FCCERS-R; $M = 4.0$ at Level 4). While the equivalence of quality scores across these three ERS scales is not supported by research evidence, the results taken together suggest the need to strengthen quality standards and assessment procedures for all types of care, so that child care providers at the highest rated levels are providing care that is at or above threshold levels recommended to impact children's developmental outcomes (Zaslow, Martinez-Beck, Tout, & Halle, 2011).

Maine: assessing the output of the rating process

As in Indiana, differences in program quality were measured using Environmental Rating Scales (ERS) mean scores at the classroom level, and these scores were not part of the QRIS standards or ratings. The results presented here are from factorial ANOVAs to examine the effects of Step Level, ERS scale type and child care program type on the dependent variable ERS mean score. Table 4 provides the adjusted mean scores for all 307 classrooms and by each program type by Step Level. Table 4 is extracted from the full report on Maine's QRIS (see Lahti et al., 2011).

The results show an overall significant difference between Step Level and ERS mean score at the classroom/setting level ($F = 5.02$; $df = 3, 307$; $p = .002$). Results of post hoc Bonferroni tests showed a significant difference between Step One and Step Four programs ($p = .001$) and between Step Two and Step Four programs ($p = .001$). The total variance of the mean ERS score explained by Step Level was only 5%, indicating weak relationships between the variables. Comparisons of the program type mean ERS quality scores indicated a difference only between the family child care home scores and the center-based scores ($p < .001$). The family child care home mean scores were lower at each Step Level than the center-based setting scores with the exception of scores at Steps Three and Four. There did not appear to be any significant differences at Step Three or Four between the center-based and Head Start type settings ($p = .97$). The results provide some evidence for differences in rated quality, with higher ERS means for higher tier or step programs, most distinctly for family child care homes. Overall these mean scores suggest the need for considerable efforts at quality improvement, considering that the majority of settings are scoring below the "5" or "good" level on the ERS measures.

Maine: parent level data on QRIS program quality

In the Maine validation study, parents in programs selected for observation were asked to complete an anonymous survey that focused on services received by the parent and the parent's perceptions of the quality of the program. The belief was that parents served by higher Step level programs should be receiving more supports and services and therefore may rate the program higher in

level of quality. The response rate over the three-year study period was approximately 26% ($N = 1478$). These results are extracted from the full report on Maine's QRIS (see Lahti et al., 2011).

Parental perception of program quality was measured by the 15-item Emlen scale, see Emlen, Koren, and Schultze (2000), and was found not to be correlated to Step Level rating (Pearson's $r = .010$, $p = .68$). In terms of services parents should have received according to program standards, a majority of parents reported not receiving: information about other government services for their child; opportunities for parent engagement with the program; daily communication from the program about their child; and being provided an up to date written parent hand-book from their provider. There did not appear to be any difference in step level in terms of parents not consistently receiving these types of services as required by the program quality standards according to parent reports. The use of these data by QRIS administrators was primarily for monitoring purposes focused on services and or supports parents should have received based on requirements in the QRIS standards. While parents were asked about perceptions of program quality, due to a strong desire to reach out to parents as a key stakeholder in the QRIS, that information was not relied upon for program planning or program improvement.

Approach 4: relate ratings to children's development

The fourth approach to validation focuses on children's development. In many respects, this is the final step in validating a QRIS, and one that arguably should be delayed until the questions raised in the earlier approaches are addressed and changes made to the system as necessary. It may even be possible that new data will emerge that makes the costly and difficult effort involved in assessing child outcomes unnecessary. For example, if studies begin to show consistently that certain inputs, e.g., ratings-based coaching lead to substantial improvements in indicators such as instructional support, and if instructional support or other indicators is found to consistently promote improved child outcomes, it may be possible to argue that the inclusion of those inputs and measures of those outputs may suffice.

The logic models that underlie QRISs typically assert that higher quality care will be associated with improved child outcomes. Therefore, one important piece of validation evidence concerns whether children make greater developmental gains in programs with higher program-level QRIS ratings than in programs with lower ratings. While a definitive evaluation of QRIS impact on child outcomes would consist of an experimental study with random assignment of providers and children to QRIS levels, ethical and practical considerations often make experiments impractical, at least on a state-level scale. Instead, current studies evaluating QRIS validity in terms of child outcomes using this approach do not attempt to evaluate causal linkages. Instead, they examine whether the QRIS ratings and the quality components that comprise the ratings are associated in expected ways to measures of children's development. Showing significant associations between QRIS-rated quality would be a first step, a necessary but not sufficient result to demonstrate causal inferences about how QRIS quality influences children's outcomes.

To date, few QRIS validation studies have incorporated children's outcomes. Maine did not include this approach to validation. As Elicker and Thornburg (2011) note, results from such studies are mixed, at least in part because of the challenges of conducting them. A primary challenge is the inability to control for all the factors that may confound the quality-outcome correlations for children whose families have selected programs in a non-random way. Additional challenges include the difficulty of recruiting of programs and children across all quality levels; lack of information about the amount of care children received in each setting (dosage); lack of appropriate outcome measures for children of diverse ages, abilities, cultures and linguistic backgrounds; and, lack of variation in the quality of participating QRIS programs. As noted above in the discussion of Approach 3, measurement error remains a problem.

Indiana: examination of ratings associated with children's outcomes

To examine validity-related questions about children's development in the context of the Indiana QRIS, the evaluators assessed the developmental status of 557 children (249 infants/toddlers; 308 preschoolers) who were in the care of QRIS providers. Two children per classroom or home were randomly selected in approximately equal numbers at all four QRIS levels. Data from parent interviews describing annual family income and parents' education levels and participation in the CCDF voucher program were used as control covariates in the analyses. The basic validity question explored was: are children in higher-rated QRIS care functioning at higher levels, socially and cognitively, than children in lower rated care?

It is important to point out that this study of quality and child outcome associations was cross-sectional, with all data collected at one point in time. As mentioned earlier, exploring these correlational relationships does not substitute for longitudinal or experimental designs that can better evaluate the causal impact of the QRIS on child outcomes. However in the implementation phase of QRIS, it is useful to explore the developmental status of participating children, how they are distributed in the child care system, and whether associations between quality measures and measures of children's functioning are occurring in the expected direction (Elicker & Thornburg, 2011).

Bivariate correlations and multiple regression models were used to explore the associations between children's development and the three measures of child care quality: QRIS ratings (4 levels); ERS global quality scores; and a CIS composite rating of positive adult-child interactions. All regression models included parent education, household income, and type of child care. No significant correlations were found between the four-level QRIS quality ratings and either infant/toddler or preschooler developmental status. Some of the researcher-observed quality measures were mildly but significantly correlated with child development measures. For preschoolers, CIS positive interactions were correlated with social competence ($r = .17^{**}$) and receptive language ability ($r = .17, p < .01$). For infants and toddlers, ERS global quality scores were associated with social competence ($r = .15, p < .01$), and total CIS positive interactions were associated with cognitive/language competence ($r = .17, p < .01^{**}$). These significant correlates were entered as predictors in regressions of child outcomes on the quality variables, controlling for the family SES variables (parent education level and household income) that were also significantly correlated with the child outcome variables. As a result, for preschoolers, CIS positive adult-child interactions significantly predicted children's receptive language ability, after controlling for family SES ($b = .12, p < .05$). For infants and toddlers, CIS positive adult-child interactions significantly predicted children's cognitive/language competence, after controlling for family SES ($b = .14, p < .05$).

Family income was also a significant predictor, $b = .23, p = .009$.

Therefore while QRIS rated levels were not significantly associated with any child development measures for either infants/toddlers or for preschoolers, ERS and CIS quality measures were moderately associated with aspects of children's development. Specifically, after controlling for family SES, it was the positive quality of interaction between adults and children that was associated with language and cognitive functioning, for both preschoolers and infants and toddlers.

Therefore it appeared that the specific aspects of child care quality assessed by the ERS and CIS measures are more likely to be associated with children's development than are the composite of quality indicators represented by the 4-level QRIS ratings. This was true even though the QRIS ratings and the ERS and CIS were significantly correlated with each other. As a result, in Indiana, further refinement of QRIS standards and procedures is taking account of these findings, especially by identifying ways to strengthen QRIS standards and ratings to include the quality of adult-child interactions.

Discussion

Limitation to validation study designs

Both of these state studies provide results that describe linear associations among variables. The study designs are limited due to the fact that the investigators have no control over how the QRIS systems are implemented which affects enrollment and therefore sample sizes and selection of measurement strategies were also not in the sole control of the investigator. It will be interesting as additional studies are done and where non-linear associations are found to determine the impact this has on outcomes. These field studies were conducted with all the limitations associated with working in a developing system with multiple stakeholders. While the design presents a limitation in terms of arguing for causality and application of more sophisticated analytic approaches, it should be noted that the state agency program managers and other stakeholders in both states found the information generated from these studies of high value in terms of system planning, program improvements, and resource allocation. Depending upon context and resources, limitations to these two study designs can be remedied in future studies by such design choices as having programs that are on a waiting list be compared to programs already participating in their state's QRIS.

Validation of QRIS is a process that needs attention over time, using more than one approach

The examples from Indiana and Maine illustrate how these validation approaches can work in practice, with tangible benefits for system improvement. These validation activities are specific to the design and implementation of each state QRIS. We believe it is important to stress to QRIS policy leaders that each of the four validation approaches needs to be used appropriately, considering the developmental stage of the QRIS and the unique features of the setting and QRIS. For example, states with QRIS in development can use the four approaches as a framework for planning how to validate their system. Developmentally then, an initial focus of QRIS design would be to validate the key concepts used in the QRIS design. The four approaches highlighted in this paper need to be considered as part of an ongoing process, not a one-time event. As states progress in their implementation of QRISs, more descriptive research is necessary to understand better how these validation approaches work in other settings, for example with point-based QRISs (rather than "building block" QRISs, like those of Indiana and Maine). Use of these approaches enables cross system comparisons,

which will allow for the identification of common threats to validity and useful strategies to enhance the validity of a state QRIS.

Validation and early care and education system constraints

A QRIS is not merely a program-level quality-improvement intervention, it is a policy lever for strengthening a state's overall early care and education system that reaches beyond child care (Schaack, Tarrant, Boller, & Tout, 2012). The two state validation efforts highlighted in this paper reflect the challenges and constraints common to other state experiences with validation activities (Lahti, Sabol, Starr, Langill, & Tout, 2013). On-site observations of global program quality and establishing and maintaining inter-rater reliability for QRIS raters is a time consuming and costly endeavor. Keen interest in school readiness may pressure program administrators to collect child outcome level data before a QRIS is well established. Current research on the measures that are in the widest use to predict child outcomes appear to do so consistently, especially for children at-risk, but with modest levels of association with program quality as measured by the ERS's (for example, Burchinal et al., 2011). We recommend to policy makers to always take into account that any validation study is occurring within a dynamic system. System-level constraints such as varying resources available to programs, different type and design of programs, and challenges to measuring quality and reliably collecting information about program quality all influence the design and implementation of state-level QRIS validation studies.

Validation research is critical for performance measurement and improvement for a state QRIS

The 2012 Child Care Development Fund (CCDF) Plan preprint for fiscal years 2014–2015 includes a much larger focus on QRIS (U.S. Department of Health & Human Services, Administration for Children and Families, 2011). In this document, a QRIS is defined as a "...systematic framework for evaluating, improving, and communicating the level of quality in early childhood programs." States are expected to provide a self-assessment based on current program quality initiatives from a set of questions that are also organized according to a "QRIS framework." Validation of program standards or assessment tools is mentioned specifically in relation to information states must provide about data and performance measures on program quality.

The information generated from QRIS validation activities can be used to inform efforts for continuous quality improvement. For example, both Indiana and Maine found that, for at least some types of providers, enrollment patterns in the QRIS, and lack of movement by programs once they are in the QRIS, is resulting in a large proportion of providers at the lower-rated quality levels of the quality tiers. System-level, quality-improvement responses to this information could be to re-assess the design of the system in terms of the ability of programs to meet standards at each tier, or to focus training/technical assistance on specific quality standards that are most challenging for providers to meet. At the same time, care should be taken to ensure that standards reflect current knowledge about the specific indicators and levels of quality most likely to produce the desired child outcomes. Findings from validation studies can be part of the information that state child care administrators use to assess the overall performance of the state early care and education system. The performance data could then be used by program administrators in making decisions about monitoring programs in a differential manner by visiting those programs more often who are having difficulty meeting QRIS standards (Fiene, 2013). We recommend that one focus of future research be learning more about whether and how information from the results of validation studies are used to improve system and program level performance in QRIS.

It will be interesting to determine as more validation studies are completed to analyze the differences between levels and how often lower quality is present in the top level which is the case with measuring compliance with licensing standards (Fiene & Nixon, 1985). Key areas to look at will be the movement of programs from one level to another, how long this takes, and are the increments equal or not in terms of quality improvement.

Another area to be explored which may have an impact on overall QRIS implementation are the fiscal constraints that many states are experiencing due to the recent recession and lower levels of federal funding. It would be interesting to note differences amongst states with large investments in quality improvement initiatives and those states with smaller investments.

It is important to remember that the QRIS is a policy lever and the validation of child care quality standards in a QRIS is a new phenomenon in early care and education policy-making. These four recommended approaches to QRIS validation, illustrated by validation research in these two states, even with their limitations, did provide policy makers and program administrators with information that guided efforts at system quality improvement. The use of these approaches in other studies will create a common nomenclature for better understanding threats to validity in a QRIS and ultimately increase our understanding of how best to design a QRIS that meets the needs of the parents, providers and children it serves.

References

- Alkon, A. (2013). *National health and safety tool*. San Francisco, CA: California Child Care Health Program.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education [AERA/APA/NCME]. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology*, 10, 541–552. [http://dx.doi.org/10.1016/0193-3973\(89\)90026-9](http://dx.doi.org/10.1016/0193-3973(89)90026-9)
- Berry, D. J., Bridges, L. J., & Zaslow, M. J. (2004). *Early childhood measures profile*. Retrieved from: <http://aspe.hhs.gov/hsp/ECMeasures04/report.pdf>
- Briggs-Gowan, M. J., & Carter, A. S. (2002). *Brief infant-toddler social and emotional assessment (BITSEA) manual, version 2.0*. New Haven, CT: Yale University.
- Bryant, D. M. (2010). *Observational measures of quality in center-based early care and education programs (OPRE Research-to-Policy, Research-to-Practice Brief 2011-10c)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Bryant, D. M., Burchinal, M., & Zaslow, M. (2011). Empirical approaches to strengthening the measurement of quality: Issues in the development and use of quality measures in research and applied settings. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), *Quality measurement in early childhood settings* (pp. 33–47). Baltimore, MD: Paul H. Brookes.
- Burchinal, M. R., Peisner-Feinberg, E. S., Bryant, D. M., & Clifford, R. (2000). Children's social and cognitive development and child-care quality. Testing for differential associations related to poverty, gender, or ethnicity. *Journal of Applied Developmental Science*, 4, 149–165. http://dx.doi.org/10.1207/S1532480XADS0403_4
- Burchinal, M. R., Vandergrift, N., Pianta, R., & Mashburn, A. J. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in prekindergarten programs. *Early Childhood Research Quarterly*, 25(2), 166–176. <http://dx.doi.org/10.1016/j.ecresq.2009.10.004>
- Burchinal, M. R., Kainz, K., Cai, Y., Tout, K., Zaslow, M., Martinez-Beck, I., et al. (2009). *Early care and education quality and child outcomes (Research-to-Policy Research-to-Practice Brief: OPRE Research-to-Policy Brief #1)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, US Department of Health and Human Services, and Child Trends.
- Burchinal, M. R., Kainz, K., & Cai, Y. (2011). How well do our measures of quality predict child outcomes? A meta-analysis and coordinated analysis of data from large scale studies of early childhood settings. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), *Quality measurement in early childhood settings* (pp. 11–27). Baltimore, MD: Paul H. Brookes.
- Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly*, 25(2), 140–165. <http://dx.doi.org/10.1016/j.ecresq.2009.11.001>
- Cizek, G. J. (2007, August). Introduction to validity. In *Presentation to the National Assessment Governing Board of NAEP* Washington, DC.
- Clifford, R. (2005). Structure and stability in the early childhood environment rating scale. In H. Schonfeld, S. O'Brien, & T. Walsh (Eds.), *Question of quality: Proceedings of a conference on defining, assessing and supporting quality in early childhood care and education* (pp. 12–21). Dublin, Ireland: Centre for Early Childhood Development and Education, St. Patrick's College.

- Cochran, M. (2007). *Finding our way: The future of American early care and education*. Washington, DC: Zero to Three.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Elicker, J., Langill, C. C., Ruprecht, K., & Kwon, K. (2007). *Paths to QUALITY™: A child care quality rating system for Indiana. What is its scientific basis?* Lafayette, IN: Purdue University, Center for Families.
- Elicker, J., & Thornburg, K. (2011). *Evaluation of quality rating and improvement systems in early childhood programs and school age care: Measuring children's development (Research to Policy, Research to Practice Brief, OPRE 2011-11c)*. Washington, DC: Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation.
- Elicker, J., Langill, C. C., Ruprecht, K., Lewsader, J., & Anderson, T. (2011). *Evaluation of "Paths to QUALITY" Indiana's child care quality rating and improvement system*. West Lafayette, IN: Purdue University Center for Families.
- Elicker, J., Langill, C., Ruprecht, K. M., Lewsader, J., Anderson, T., & Brizzi, M. (2013). Paths to QUALITY: Collaborative evaluation of a new child care quality rating and improvement system. *Early Education and Development*, 24(1), 42–62. <http://dx.doi.org/10.1080/10409289.2013.736127>
- Emlen, A., Koren, P., & Schultze, K. (2000). *A packet of scales for measuring quality from a parent's view*. Portland, OR: Portland State University, Regional Research Institute for Human Services, The Oregon Child Care Research Partnership.
- Fiene, R. (2013). *Differential monitoring logic model and algorithm (DMLMA): A new early childhood program quality indicator model (ECPQIM) for early care and education regulatory agencies*. Middletown, PA: Research Institute for Key Indicators.
- Fiene, R., & Nixon, M. (1985). Instrument based program monitoring and key indicators in child care. *Child Care Quarterly*, 14(3), 198–214.
- Fuller, B., Loeb, S., Kagan, S. L., & Carrol, D. (2004). Child care in poor communities: Early learning effects of type. *Quality, and Stability, Child Development*, 75(1), 47–65.
- Harms, T., Clifford, R., & Cryer, D. (2005). *Early childhood environmental rating scale-revised*. New York, NY: Columbia University Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. (2006). *Infant toddler environmental rating scale-revised*. New York, NY: Columbia University Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. (2007). *Family day care environmental rating scale-revised*. New York, NY: Columbia University Teachers College Press.
- Howes, C., Whitebook, M., & Phillips, D. (1992). Results on the National Child Care Staffing Study concerning effectiveness and characteristics of teachers across the United States. *Child & Youth Care Forum*, 21(6), 399–414.
- Howes, C., Burchinal, M., Pianta, R. C., Bryant, D. M., Early, D. M., Clifford, R. M., et al. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Child & Youth Care Forum*, 37(1), 27–50. <http://dx.doi.org/10.1016/j.ecresq.2007.05.002>
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Education Measurement*, 38, 319–342.
- Karoly, L. A., Zellman, G. L., & Perlman, M. (2013). Understanding variation in classroom quality within early childhood centers: Evidence from Colorado's quality rating and improvement system. *Early Childhood Research Quarterly*, 28, 645–657. <http://dx.doi.org/10.1016/j.ecresq.2013.05.001>
- LaFreniere, P. J., & Dumas, J. E. (1997). Social competence and behavior evaluation in children ages 3 to 6 years: The short form (SCBE-30). *Psychological Assessment*, 8(4), 369–377. <http://dx.doi.org/10.1037/1040-3590.8.4.369>
- Lahti, M., Cobo-Lewis, A., Dean, A., Rawlings, S., Sawyer, E., & Zollitsch, B. (2011). *Maine's quality for ME – child care quality rating and improvement system (QRIS): Final evaluation report*. Portland, ME: Cutler Institute for Health and Social Policy, Muskie School of Public Service, University of Southern Maine.
- Lahti, M., Sabol, T., Starr, R., Langill, C., & Tout, K. (2013). *Validation of quality rating and improvement systems (QRIS): Examples from four states (Research-to-Policy, Research-to-Practice Brief OPRE 2013-036)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Lugo-Gil, J., Sattar, S., Boss, C., Boller, K., Tout, K., & Kirby, G. (2011). *The quality rating and improvement system (QRIS) evaluation toolkit (OPRE Report #2011-31)*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation.
- Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service, Inc.
- NICHD Early Child Care Research NICHD Network. (2000). Characteristics and quality of child care for toddlers and preschoolers. *Applied Developmental Sciences*, 4, 116–135.
- NICHD Early Child Care Research NICHD Network. (2002). Child care structure → process → outcome: Direct and indirect effects of child care quality on young children's development. *Psychological Science*, 13, 199–206.
- Maine, Department of Health and Human Services. (2008). *Quality for ME: Child care quality rating system program standards*. Retrieved from: <http://www.maine.gov/dhhs/ocfs/ec/occhs/quality.cbc.pdf>
- Mathematica Policy Research. (2011). *Defining and measuring quality: An in-depth study of five child care quality rating and improvement systems (OPRE Report 2011-29)*. Retrieved from: <http://www.whitehouse.gov/sites/default/files/omb/budget/fy2011/assets/management.pdf>
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system manual, Pre-K*. Baltimore, MD: Paul H. Brookes.
- QRIS National Learning Network (2014). Retrieved from: <http://qrisnetwork.org/qris-state-contacts-map>
- Schaack, D., Tarrant, K., Boller, K., & Tout, K. (2012). *Quality rating and improvement systems, frameworks for early care and education systems change*. In S. L. Kagan, & K. Kauerz (Eds.), *Early childhood systems, transforming learning* (pp. 71–86). New York, NY: Teachers College Press.
- Smith, S., Robbins, T., Stagman, S., & Kreader, J. L. (2012). *Practices for promoting young children's learning in QRIS standards*. Retrieved from: <http://www.nccp.org/publications/pdf/text.1070.pdf>
- The Concept System® (2012). BT computer software. PN Concept Systems, Incorporated, PL Ithaca, NY.
- Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). *Issues for the next decade of quality rating and improvement systems, OPRE Issue Brief*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G., & Boller, K. (2010). *Compendium of quality ratings systems and evaluations*. Washington, DC: Child Trends and Mathematica Policy Research, for Office of Planning, Research and Evaluation, Administration for Children and Families, US Department of Health and Human Services.
- U.S. Department of Health and Human Services, Administration for Children and Families. (2011). *Program instruction (Child Care and Development Block Grant Act of 1990)*. Retrieved from: <http://www.acf.hhs.gov/sites/default/files/occ/pi2011.03.pdf>
- Zaslow, M., Martinez-Beck, I., Tout, K., & Halle, T. (2011). *Quality measurement in early childhood settings*. Baltimore, MD: Paul H. Brookes.
- Zellman, G. L., & Fiene, R. (2012). *Validation of quality rating and improvement systems for early care and education and school-age care (Research-to-Policy, Research-to-Practice Brief OPRE 2012-29)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Zellman, G. L., Perlman, M., Le, V., & Setodji, C. M. (2008). *Assessing the validity of the Qualistar Early Learning quality rating and improvement system as a tool for improving child-care quality (MG-650-QEL)*. Santa Monica, CA: RAND Corporation.
- Zellman, G. L., Brandon, R. N., Boller, K., & Kreader, J. L. (2011). *Effective evaluation of quality rating and improvement systems for early care and education and school-age care (Research-to-Policy, Research-to-Practice Brief OPRE 2011-11a)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Zellman, G. L., & Karoly, L. A. (2012). *Incorporating child assessments into state early childhood quality improvement initiatives*. Santa Monica, CA: RAND Corporation.

Child Care Health Consultation Improves Infant and Toddler Care

Rosemary Johnston, RN, BSN, MSN, Beth A. DelConte, MD, FAAP, Libby Ungvary, MEd, Richard Fiene, PhD, & Susan S. Aronson, MD, FAAP

ABSTRACT

Introduction: Many families enroll their infants and toddlers in early education and child care programs. The Pennsylvania Chapter of the American Academy of Pediatrics recruited 32 child care centers that care for infants and toddlers to be linked with a child care health consultant (CCHC).

Method: Project staff assigned the centers alternately to an immediate intervention or a 1-year delayed intervention (contrast) group. At entry into the project, and then 1 and 2 years later, an evaluator assessed center compliance with 13 standards for infants and toddler care selected from *Caring for Our Children: National Health and Safety Performance Standards* (3rd ed.). Project staff linked the Immediate Intervention centers with a CCHC in Year 1. In Year 2, in a crossover comparison, project staff linked Contrast centers with a CCHC.

Results: Working with a CCHC effectively improved compliance with some selected health and safety standards. *J Pediatr Health Care.* (2017) ■, ■-■.

KEY WORDS

Child care, child care health consultation, health and safety, infants and toddlers

INTRODUCTION

Nationally, about 48% of children younger than 3 years of age are enrolled in organized child care facilities (Laughlin, 2013). Early educators (child care staff) care for these children for many hours and many days. The quality of their care has lifelong impact on their physical, developmental, and social-emotional well-being (Garcia, Heckman, Leaf, & Padros, 2016).

In 2013, the Early Childhood Education Linkage System (ECELS), a program of the Pennsylvania (PA) Chapter of the American Academy of Pediatrics (AAP)

Rosemary Johnston, Infant Toddler Quality Improvement Project Coordinator, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA.

Beth A. DelConte, Pediatric Advisor, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA.

Libby Ungvary, Director, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA.

Richard Fiene, Research Psychologist, Research Institute for Key Indicators, Middletown, PA, and Affiliate Professor, Prevention Research Center, The Pennsylvania State University, University Park, PA.

Susan S. Aronson, Pediatric Advisor, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA, and Retired Clinical Professor of Pediatrics, The University of Pennsylvania and The Children's Hospital of Philadelphia, Philadelphia, PA.

This project was supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and

Human Services (HHS) under grant No. H25MC26235 Community-Based Integrated Service Systems. The information, content, and/or conclusions in this paper are those of the author and should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS, or the U.S. Government.

Conflicts of interest: None to report.

Correspondence: Rosemary Johnston, RN, BSN, MSN, Infant Toddler Quality Improvement Project Coordinator, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System (ECELS), 1400 North Providence Rd., Ste. 3007, Rose Tree Corporate Center II, Media, PA 19063; e-mail: rljrmj@msn.com.

0891-5245/\$36.00

Copyright © 2017 by the National Association of Pediatric Nurse Practitioners. Published by Elsevier Inc. All rights reserved.

<http://dx.doi.org/10.1016/j.pedhc.2017.05.005>

received a 3-year grant from the Maternal and Child Health Bureau (MCHB). The purpose of the grant was to “improve state infant/toddler [I/T] child care quality initiatives (Quality Rating and Improvement Systems [QRIS] and professional development)...” MCHB’s grant required selection and promotion of 10 or more standards from a list provided by MCHB from *Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs*, 3rd ed. (CFOC3; AAP, American Public Health Association, & National Resource Center for Health and Safety in Child Care and Early Education, 2011).

Child care programs in PA’s QRIS, called Keystone STARS, are ranked from the entry level at STAR 1 to STAR 4. To earn a rating, programs must comply with state regulations and meet the requirements listed for the designated STAR level on the PA Key Web site (www.pakeys.org). For a STAR 4 rating, a center that serves infants and toddlers must have scores at or above 5 (*good*) on the seven subscales of the Infant and Toddler Environment Rating Scale–Revised Edition (ITERS-R; Harms, Cryer, & Clifford, 2006). The Personal Care Routines subscale of the ITERS-R has some health and safety items. Scores in this subscale and on health and safety items in some of the other subscales are among the lowest scoring ITERS-R items in PA and elsewhere. This finding is reported by the PA Key Program Quality Assessment Team (2016) and by the authors of the ITERS-R (Harms and Cryer, personal communication, 2014).

Child care health consultants (CCHCs) use observation, education, collaborative decision making, coaching, and mentoring to achieve quality improvement in the QRIS (Zaslow, Tout, & Halle, 2012). CCHCs base their work on needs and feasible implementation. For more than a decade, published research has confirmed that child care health consultation is an effective approach to improving health and safety compliance with national child care standards (Alkon & Bernzweig, 2008; Alkon et al., 2008; Alkon, Bernzweig, Kim, Wolff, & Mackie, 2009; Alkon et al., 2014; Alkon et al., n.d.; Alkon, Sokal-Gutierrez, & Wolf, 2002; Banghart & Kraeder, 2012; Carabin et al., 1999; Crowley, 2006; Isbell et al., 2013; Moon & Oden, 2005; Organizational Research Services & Geo Education and Research, 2007; Pacific Research and Evaluation, 2007, 2008; Ramler, Nakatsukasa-Ono, Loe, & Harris, 2006; Roberts et al., 2000a, 2000b). Most of these studies did not specifically target care for infants and toddlers.

Published studies document the following specific improvements associated with involvement of a CCHC. Sanitation and hygiene reduced respiratory and gastrointestinal illness and days absent for illness among young children in group care (Carabin et al.,

1999; Kotch et al., 2007; Roberts et al., 2000a, 2000b). Nationally recommended practices related to active play, nutrition, and food handling were adopted (Alkon et al., 2014). Policies and procedures accompanied by staff training reduced hazards and injuries (Kotch, 2002; Organizational Research Services & Geo Education and Research, 2007). Training about safe infant sleep positioning and the infant sleep environment reduced risk of sudden infant death syndrome (Moon & Oden, 2005). Better monitoring and tracking of immunization data in child care programs was associated with more children having up-to-date vaccine documentation (Alkon & Bernzweig, 2008).

The PA AAP established ECELS in 1989. ECELS maintains a CCHC Registry and regularly communicates with registered CCHCs to provide professional development, technical assistance, and tools to enable their implementation of the CCHC role. PA’s CCHCs include private and public health service providers and health professionals who teach in academic settings. Funding for CCHC work is unpredictable, making recruitment, education, and retention of CCHCs challenging.

PA’s child care regulations require that child care providers have documents showing that enrolled children are up to date with preventive health services recommended by the AAP, including “a review of the child’s immunized status according to recommendations of the ACIP [Advisory Committee on Immunization Practices]” (PA Department of Human Services, 2008). This regulation is not enforced. Few providers use any reliable way to ensure that enrolled children are up to date. ECELS encourages child care centers to use a well-tested and routinely updated online software application called WellCareTracker™ (Weinburg, 2002) to check child health records for up-to-date routine preventive health services. It is described, demonstrated, and offered for subscription at www.wellcaretracker.org. Using WellCareTracker™ eases the burden for child care providers to comply with the regulation and remind families to obtain these services in a timely manner.

METHODS

Design

The PA AAP’s MCHB-funded Infant-Toddler Quality Improvement Project (ITQIP) was conducted by ECELS using a randomly assigned clinical trial with a crossover comparison of centers assigned to an immediate intervention or delayed intervention (comparison) group. ECELS (a) assessed child care center practices related to I/T care for 13 selected CFOC3 standards (AAP et al., 2011) and (b) assessed whether compliance with these practices improved when centers were linked with a CCHC.

Selection of the CFOC3 standards addressed in ITQIP

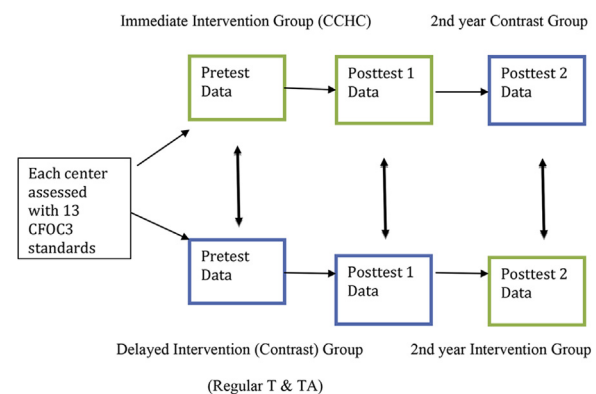
With input from early care and education stakeholders, ECELS chose 13 CFOC3 standards (AAP et al., 2011) from a list provided by MCHB (Box 1). The selection criteria were that the standard is (a) associated with the highest and most common risks of harm to I/T (AAP, American Public Health Association, & National Resource Center for Health and Safety in Child Care and Early Education, 2013), (b) measurable and amenable to improvement with technical assistance and professional development provided by a CCHC over a 12-month period, and (c) found by state inspectors to have a high level of noncompliance according to state data (PA Office of Child Development and Early Learning, 2010).

Evaluation plan

The evaluation plan is a classic randomly assigned crossover clinical trial. See Figure 1 for the evaluation plan logic model.

The ITQIP staff and consultants developed the evaluation tool described below. The ITQIP Project Coordinator (first author) and the evaluators collected data from participating centers at three points: when centers enrolled in the study (Pretest) and then 1 year (Posttest 1) and 2 years later (Posttest 2). One of the consultants (fourth author) compared the two groups on the pretest for equivalency and then on each of the two posttests. These data are discussed in the Results: Immediate Intervention Versus Delayed Intervention (Contrast) Group section. One

FIGURE 1. Evaluation plan logic model.
CCHC, child health care consultant; T, training; TA, technical assistance.



year after the pretest data were collected, the participating centers were switched to a crossover comparison format. At this point, ITQIP ended the subsidy for the CCHCs who were working with the centers in the immediate intervention group and provided the subsidized CCHC linkage to the centers in the delayed intervention (contrast) group.

When a center enrolled in ITQIP, the ITQIP coordinator interviewed the center director by phone. She gathered demographic data, including the number of enrolled I/Ts, where and when I/T activities occurred in the center, and the number of children who met the MCHB definition of special health needs. She asked the director to submit up to five of any care plans the center had for these children, redacted for confidentiality. The MCHB definition of a child with special health care needs is noted in CFOC3 standard 3.5.0.1 as “a child who has or is at increased risk for chronic physical, developmental, behavioral or emotional conditions and who requires health and related services of a type or amount beyond that required by children generally” (AAP et al., 2011).

The ITQIP coordinator selected the rooms for the evaluator to observe as those with the largest number of children in the age group. The evaluators recorded observations in one infant and one toddler room at each center.

The evaluator collected a random sample of immunization records for up to 10 infants and 10 toddlers with the names redacted for confidentiality. The ITQIP coordinator used WellCareTracker™ software to check these immunization records. The ITQIP coordinator evaluated the care plans that the director submitted for the presence of the appropriate components from the list of the 14 components specified in CFOC3 standard 3.5.0.1. (AAP et al., 2011) and a 15th component, the presence of the health care provider’s signature, that is required by PA regulations (Box 2).

BOX 1. CFOC3 standards chosen for ITQIP

- 1.4.5.2 Child Abuse and Neglect Education
 - 3.4.4.1 Recognizing and Reporting Suspected Child Abuse, Neglect, and Exploitation
 - 2.1.2.1 Personal Caregiver/Teacher Relationships for Infants and Toddlers
 - 2.2.0.2 Limiting Infant/Toddler Time in Crib, High Chair, Car Seat, and other restraining equipment
 - 3.1.3.1 Active Opportunities for Physical Activity
 - 3.1.4.1 Safe Sleep Practices and SIDS Risk Reduction
 - 3.2.1.4 Diaper Changing Procedure
 - 3.2.2.1 Situations That Require Hand Hygiene
 - 3.2.2.2 Handwashing Procedure
 - 3.6.3.3 Training of Caregivers/Teachers to Administer Medication
 - 3.5.0.1 Care Plan for Children with Special Health Care Needs
 - 5.4.5.2 Cribs
 - 7.2.0.1 Immunization Documentation
- Note. CFOC3, *Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs* (3rd ed.); ITQIP, *Infant-Toddler Quality Improvement Project*; SIDS, *sudden infant death syndrome*.

The ITQIP coordinator scored the evaluator's observations of diapering, hand hygiene, and medication administration. She promptly prepared a summary of all the findings for the center and sent the summary to the center director and the linked CCHC before the first CCHC site visit. The summary delineated areas of strengths and areas to improve based on the evaluation tool results. To facilitate use of the data by the center staff and CCHCs, the summary included the text of the evaluation tool item, the center's score on the item, and the reason why the center met or did not meet the standard. The CCHC contacted the center within 2 weeks after receiving the summary to set up the initial site visit.

Evaluation Tool

The ITQIP staff prepared the items on the evaluation tool from performance guidelines specified in the 13 selected *CFOC3* standards (AAP et al., 2011). ITQIP consultants (fourth and fifth authors) and the ECELS staff reviewed the tool for clarity and validity of content. After several rounds of revisions, the ITQIP coordinator and a prospective ITQIP evaluator field-tested the tool, further revised it, and then field-tested it again, this time

testing for interrater reliability with two evaluators independently and simultaneously using the tool.

The ITQIP evaluation tool has four sections: (a) Demographic Information collected in the phone interview (35 items), (b) Observations (64 items), (c) Interview Questions (28 items), and (d) Documents (14 items). The score awarded to items on the evaluation tool was based on the criteria listed in Box 3. A score of 2 or 3 for an item was considered a strength, and a score of 0 or 1 for an item was considered an area to improve. This total score was the sum of the scores for each item. The total number of scorable items on the evaluation tool is 106, with a maximum score of 318. The documents assessed include training records, written policies, care plans for children with special needs, immunization data, and PA child abuse clearances.

ITQIP assigned each scorable item to one of the 10 topic areas addressed by the 13 *CFOC3* standards selected for the project (AAP et al., 2011). See Table 1.

Sampling design: Recruitment, random assignment, and retention of centers

ECELS recruited Keystone STAR 2 and STAR 3 centers by distributing a flyer about the project. Programs with higher STARS ratings qualify for higher payments for children whose care is state subsidized. The highest payments are for children enrolled in STAR 4 centers. The increased payment for a higher rating is a quality improvement incentive. Also, ECELS offered participating centers three free \$10 credit-awarding reviews for ECELS self-learning modules. The flyer was included in the newsletters of a variety of organizations: four of the five regional state-supported sources of professional development (Regional Keys), the PA Child Care Association, the Pittsburgh Association for the Education of Young Children, and United Way. Because the northwestern region of the state has the fewest centers, recruitment from that region was not attempted.

As the centers joined ITQIP, the project coordinator assigned them alternately to one of two groups, either the immediate intervention group or the delayed intervention (contrast) group. ITQIP enrolled centers from all four targeted regions of the state.

BOX 2. Care plan components evaluated for children with special needs

1. A list of the child's diagnoses
2. Contact information for the child's health care provider and any subspecialists
3. Medications to be administered on a scheduled basis
4. Medications to be administered in an emergency with clearly stated signs and symptoms in lay language about when to give the medication
5. Procedures to be performed while in care
6. Allergies
7. Diet modification that the child requires
8. Activity modifications
9. Environmental modifications
10. Triggers that cause a reaction to avoid
11. Symptoms for caregivers/teachers to observe
12. Behavioral modifications beyond those needed for a typically developing child
13. Emergency response plans for a facility emergency and if the child has an emergency event
14. Special skills training and education required and provided for the staff
15. Health care provider signature (required by Pennsylvania regulation)

Note. Fourteen components specified in the Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs (3rd ed.) standard 3.5.0.1. (American Academy of Pediatrics et al., 2011) and a 15th required by Pennsylvania child care regulation.

BOX 3. Criteria for scores assigned to items on the evaluation tool

- 0 = Never meets item
- 1 = Partly (<50%) meets item
- 2 = Mostly (≥50%) meets item
- 3 = Fully (100%) meets the item
- NA = Not Applicable
- NOp = Not observed or no opportunity to obtain data
- DK = Don't know (interviewee response)

TABLE 1. Topic areas and number of items to score per topic

Abbreviation	Topic areas	Number of items to score per topic ^a
CA	Preventing Child Abuse	13
PR	Personal Relationships	9
LA	Limited Physical Activity of Infants	3
AO	Active Opportunity for Physical Activity	22
SS	Safe Sleep Practices/SIDS Prevention	19
MA	Medication Administration	8
DC	Diaper Changing Procedure	16
HH	Hand Hygiene	8
IM	Immunization Documentation	3
SN	Care Plans for Children With Special Needs	5

Note. SIDS, sudden infant death syndrome.

^aSee the narrative for an item-by-item explanation of those items with significance levels (p values) based on the t tests performed on each item.

Centers enrolled in ITQIP agreed to

- allow a 4- to 5-hour site evaluation once a year for 3 years,
- work with a CCHC for a period of 1 year to improve I/T health and safety,
- accept random assignment to one of the two project groups,
- provide access to redacted immunization records and care plans for evaluation,
- pay \$240.00 of the \$500 honorarium ITQIP paid to their CCHC, and
- remain in ITQIP for 3 years.

Recruitment and roles of evaluators and CCHCs Evaluators.

ITQIP recruited 17 evaluators from the list of CCHCs who had previously received CCHC training from ECELS and from the nurses in the Maternal Infant and Early Childhood Home Visiting Program. All evaluators were health professionals with pediatric experience related to observed items. Most had experience working with *CFOC3* standards (AAP et al., 2011). The evaluators learned how to use the evaluation tool by participating in a live Webinar or by using the recording of the Webinar. All evaluators received a copy of the evaluation tool and a training manual with instructions for completing the evaluation. Seven evaluators were also CCHCs in this project. None of the evaluators who were CCHCs in ITQIP were linked with centers they evaluated.

The evaluators gave their completed evaluation tools to the ITQIP coordinator to score and summarize. The coordinator reviewed each submitted evaluation

tool and then discussed the documentation with the evaluator by phone to make sure the scoring was as intended.

Child Care Health Consultants.

ECELS recruited 14 registered nurses and one physician as CCHCs. The ITQIP coordinator (first author) has worked as a CCHC for more than 15 years. She and the project's director and primary investigator, a pediatrician (second author) educated, coached, mentored, and supported the work of the CCHCs. The CCHCs participated in a Webinar about the project scope and the use of the selected *CFOC3* standards (AAP et al., 2011). They received a training manual that included the 13 selected *CFOC3* standards (AAP et al., 2011) and resources to support best practice in each of the 10 topic areas. ITQIP provided additional resources and periodic *CFOC3* updates (AAP et al., 2011).

During the site visit, the CCHC compared her observations with those in the summary and solicited concerns about health and safety practices from the center's staff. Then the director, program staff, and CCHC chose three of the 10 topics as the primary focus of the center's improvement. The CCHC helped the center staff prepare an action plan to work on the three topic areas they chose.

Action plans included filling gaps in knowledge, developing policies for staff and family handbooks, and improving staff practices. The CCHCs and center directors arranged all subsequent contacts and visits over the next 12 months.

The CCHC helped the center staff prepare an action plan to work on the three topic areas they chose.

Quarterly, the CCHCs sent the ITQIP coordinator documentation of their work and progress toward goals. The CCHCs submitted the center's initial action plan and a final action plan at the end of the year that showed what the center accomplished. ITQIP paid \$250 to the CCHCs upon receipt of the center's initial action plan and date of the first CCHC visit. ITQIP paid the CCHCs an additional \$250 after they submitted the final action plan from their 12-month linkage. Throughout the project, the ITQIP coordinator reviewed quarterly encounter forms that the CCHCs submitted to describe their work with the centers. This enabled the ITQIP coordinator to suggest ways to promote progress on action plans, including use of relevant health and safety resources.

RESULTS

Descriptive Report

ITQIP linked CCHCs with 32 centers. Of these, 16 centers were in the immediate CCHC-linked group, and 16 were in the delayed CCHC-linked group. In all,

59 directors, 348 I/T teachers and 1,490 infants and toddlers were directly involved in ITQIP. Three centers from each group dropped out, leaving 13 centers in each group at the completion of the project (Table 2).

Over the 1-year period of CCHC linkage, 12 of the 32 programs had turnovers of two to four directors. This change in center leadership made the CCHCs' work to improve I/T care very difficult. For the immediate intervention group, three of the original 16 centers withdrew from the project. One center in the delayed intervention (contrast) group closed during the project period; two others withdrew from ITQIP. Some centers dropped out because they were so overwhelmed with maintaining ratios in classrooms and staffing issues that they believed they could not focus on their action plans.

This report compares pretest, Posttest 1 and Posttest 2 scores for the 13 immediate intervention sites and 13 delayed intervention (contrast) sites that remained enrolled in ITQIP for the full 3 years.

ITQIP did not require a specific time spent in the CCHC role for each linkage. The CCHCs in the immediate intervention group provided an average of 14 hours of consultation per site (range = 2.25–28.75 hours). The CCHCs in the delayed intervention (contrast) group provided an average of 12.5 hours of consultation per site (range = 2–32 hours). The CCHCs completed quarterly encounter forms to report the total hours of services to their linked center, including a checklist of onsite, phone, and e-mail services. The most common CCHC interactions with centers included providing health education for the director and staff, onsite consultation at the facility, technical assistance by phone or e-mail, providing print or audiovisual materials, helping the facility comply with state regulations, and developing health policies and procedures.

Topics chosen by the centers in the immediate intervention group and the delayed intervention (contrast) group and the number of centers that chose each topic are shown in Table 3.

Quantitative Comparison of Evaluation Tool Scores on the Pretest Versus the Two Posttests

The scores used in the quantitative comparisons are the sum of all scores on the Evaluation Tool, not only those

for the topics that the center chose for special focus (Table 4).

Immediate intervention group

On the pretest, the range in scores was 175 to 267, with an average score of 212 out of a possible 318 points (66%). On Posttest 1, the range in scores was 213 to 297, with an average score of 254 out of a possible 318 points (79%). This change from the pretest to Posttest 1 was statistically significant ($t = -4.62, p < .0001$). Posttest 2 did not show any significant change from the average score on Posttest 1, showing that the initial results from the intervention were sustained in the next year (254 to 254).

Delayed intervention (contrast) group

On the pretest, the range in scores was 164 to 271, with an average score of 218 out of a possible 318 points (68%). On Posttest 1, the range in scores was 149 to 257, with an average score of 221 out of a possible 318 points (69%). These changes from the pretest to Posttest 1 were not significant. Posttest 2 showed significant change in the average score from Posttest 1 (221 points) to Posttest 2 (243 points; $t = -1.80, p < .08$) a year after this delayed intervention (contrast) group had received their CCHC linkage.

Immediate Intervention Versus Delayed Intervention (Contrast) Groups

The comparison of the average scores between the Immediate Intervention (212) and Delayed Intervention (Contrast, 218) groups on the pretest was not significant, showing that the groups were equivalent. The difference between the average scores of the immediate intervention (254) and delayed intervention (contrast, 221) groups on Posttest 1 was statistically significant ($t = -3.46, p < .002$), showing the effectiveness of the CCHC intervention for the immediate intervention group. Posttest 2 showed no significant difference between the change in the average postintervention scores for the immediate intervention group 12 months after their CCHC-subsidized linkage and the delayed intervention (contrast) group (254 vs. 243) at the end of their 12 months of CCHC-subsidized linkage. See Figure 2 for the crossover comparison results.

TABLE 2. Location and retention of recruited centers

Region of Pennsylvania	Immediate intervention group			Delayed intervention group		
	Centers recruited	Centers dropped out	Centers completed	Centers recruited	Centers dropped out	Centers completed
Southwest Region (Pittsburgh metropolitan area)	1	0	1	3	1	2
South Central Region (Harrisburg metropolitan area)	4	1	3	2	1	1
Northeast Region (Allentown/Bethlehem/Scranton)	3	0	3	4	0	4
Southeast Region (Philadelphia metropolitan area)	8	2	6	7	1	6
Total	16	3	13	16	3	13

TABLE 3. CFOC3 topics chosen by centers by intervention group

CFOC3 topics	Number of centers in immediate intervention group that chose each topic	Number of centers in delayed intervention (contrast) group that chose each topic
Safe Sleep Practice	11	11
Medication Administration	10	6
Child Abuse Prevention	6	1
Care Plans for Children with Special Needs	5	8
Diaper Changing Procedure	4	4
Limited Physical Activity of Infants	2	1
Hand Hygiene	2	5
Immunization	1	0
Personal Relationships	0	1
Active Opportunity for Physical Activity	0	4

Note. CFOC3, *Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs* (3rd ed.).

The crossover comparison results (Figure 2) show the relationship between the immediate intervention and the delayed intervention (contrast) groups in a crossover design. It clearly shows how effective the intervention (pretest to Posttest 1) was for the immediate intervention group and that the effects persisted after 1 year without a subsidized CCHC linkage (Posttest 1 to Posttest 2). It also shows that the intervention was effective when the delayed intervention (contrast) group was switched to receive the CCHC intervention with targeted training, technical assistance, and collaborative consultation a year after their pretest assessment (Posttest 1 to Posttest 2).

For the Immediate Intervention Group After 1 Year of Linkage With a CCHC

Among the items in each topic area (Table 1), the following items showed statistically significant improvement (pretest to Posttest 1).

Medication administration

The director had documentation that the staff who are authorized to give medications have received medica-

tion administration training within the year from a health professional ($p < .001$).

Safe sleep

The number of written safe sleep policies containing the required elements increased ($p < .05$). Teachers ($p < .01$) and parents ($p < .05$) reviewed the safe sleep policies and were educated about safe sleep practices ($p < .05$).

Child abuse

Child abuse policies contained the required elements ($p < .05$). Both infant and toddler teachers were educated about child abuse and how, as mandated reporters, they are required to personally report incidents they suspect might involve child maltreatment ($p < .001$). The number of centers having required clearance documents on file for teachers increased ($p < .05$).

Active opportunities for physical activity

Infants (birth through 12 months of age) were taken outside two to three times per day, as tolerated ($p < .05$). Toddlers (12 months through 3 years)

TABLE 4. Quantitative results of the evaluation from the pretest to two posttests

	Intervention group				Delayed intervention (contrast) group			
	Range	Average	%	Possible total	Range	Average	%	Possible total
Pretest	175–267	212 ^a	66	318	164–271	218	68	318
Posttest 1	213–297	254 ^{a,c}	79	318	149–257	221 ^{b,c}	69	318
Posttest 2	137–286	254	79	318	170–283	243 ^b	76	318

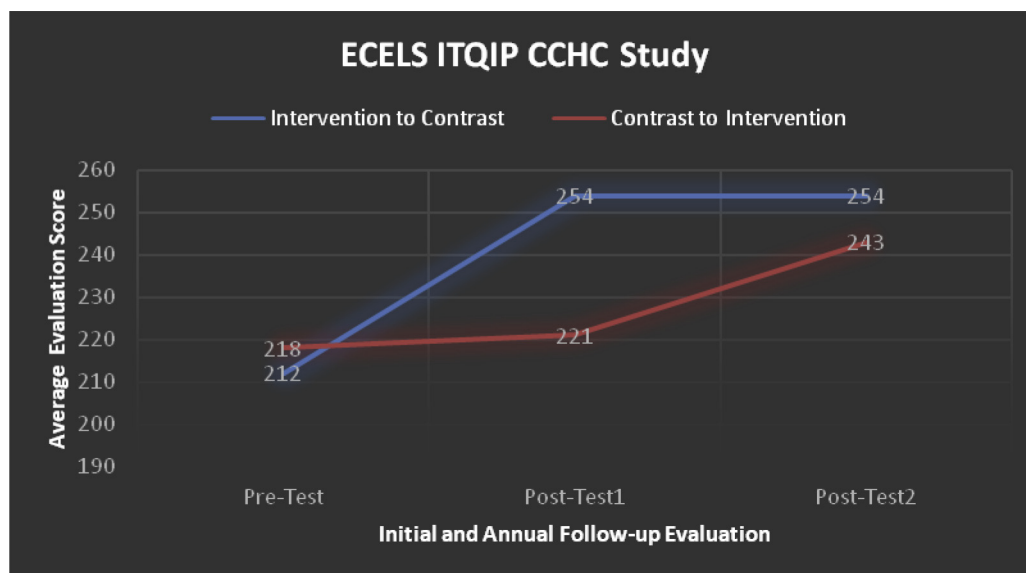
Note. CCHC, child care health consultant.

^aStatistically significant change ($t = -4.62$, $p \leq .0001$) from pretest to Posttest 1 for the immediate intervention group after the intervention of a 1-year linkage with a CCHC.

^bStatistically significant change ($t = -1.80$, $p \leq .08$) from Posttest 1 to Posttest 2 for the delayed intervention group after the intervention of 1 year of CCHC linkage.

^cStatistically significant change ($t = -3.46$; $p < .002$) for Posttest 1 between the immediate intervention group and the delayed intervention (contrast) group.

FIGURE 2. Crossover comparison results. CCHC, child care health consultant; ECELS, Early Childhood Education Linkage System; ITQIP, Infant-Toddler Quality Improvement Project.



went outside except in weather that poses a significant health risk ($p < .05$).

Diaper changing

Before the beginning of the diaper change, changing table paper was placed over the diapering surface, followed by the gathering of supplies needed for the diaper change from the containers in which they are stored and use of gloves ($p < .05$).

Hand hygiene

Observed times when toddlers ($p < .01$) and the toddler teachers/caregivers ($p < .05$) should have washed their hands showed statistically significant improvement after CCHC linkage.

For the Delayed Intervention (Contrast) Group After 1 Year of Linkage With a CCHC

Among the items in each topic area (Table 1), the following items showed statistically significant improvement (Posttest 1 to Posttest 2).

Safe sleep

Safe sleep policies that contained all the elements that should be in a safe sleep policy per *CFOC3* standard 3.1.4.1. ($p < .05$; AAP et al., 2011). The facility had documentation that parents reviewed the center's safe sleep policy and were educated about safe sleep practices ($p < .05$). There was no soft or loose bedding or other objects in a crib when an infant was in the crib ($p < .05$). Caregivers and teachers checked on sleeping infants often enough (about

every 5 minutes) to be sure that the infant was still breathing ($p < .05$).

Medication administration

The name of a child to receive medication was verified before the medication was administered to that child ($p < .05$).

Diaper changing

Bottom clothing was removed, including shoes and socks, if feet were unlikely to be kept from contacting soiled skin or surfaces. If clothing was soiled, it was removed and placed in a plastic bag ($p < .05$).

Special needs

The number of care plans submitted that included the required elements in a care plan for children with special needs per the *CFOC3* standard 3.5.0.1 increased ($p < .05$; AAP et al., 2011).

Additional Findings of Interest

Immunization documentation

Only one center chose to work on documentation of up-to-date immunization status as an action plan focus. Overall, the immunization data for the two groups showed low compliance with *CFOC3* standard 7.2.0.1 (AAP et al., 2011) and PA's immunization regulations (PA Department of Human Services, 2008). On the pretest, in the immediate intervention centers, 22% of the immunization records for infants and 43% of the immunization records for toddlers were up to date.

Little change occurred for this group on Posttest 1 (36% for infants, 43% for toddlers.) On the pretest for the delayed intervention (contrast) centers, 25% of the immunization records for infants and 40% of the immunizations records for toddlers were up to date. On Posttest 1 the delayed intervention (contrast) centers improved from 25% to 38% for infants but dropped from 40% to 27% of the records for toddlers showing up-to-date vaccines.

Care plans for children with special needs

The data for the two groups showed low compliance with *CFOC3* standard 3.5.0.1 (AAP et al., 2011) that lists the components for care plans. Combining the immediate intervention and delayed intervention (contrast) center findings for this topic, the pretest showed that 66 I/Ts were identified with special health care needs in the 32 centers initially enrolled in ITQIP. Only 15 (23%) of I/Ts with identified special health care needs had any care plan signed by a health care professional. Only 1 of 66 I/Ts with special health care needs had a care plan signed by a health care professional that had all necessary components for optimal daily and/or emergency care. Posttest 2 showed that 39 I/Ts were identified with a special health care need in the remaining 26 centers. For children identified by the centers as having a special health care need, 62% did not have a care plan. Fifteen (38%) of those with identified special health care needs had a care plan signed by a health professional. Four of the 15 care plans had all the required elements. Examples of children who had special needs and had no care plan signed by a health care provider included children with gastroesophageal reflux taking Ranitidine, febrile seizures, asthma, multiple epinephrine autoinjectors onsite, autism, nonfebrile seizures, and torticollis and plagiocephaly, which required that the child wear a helmet each day.

DISCUSSION AND CONCLUSIONS

Quality early education and child care have been shown to be associated with lifelong benefits (Garcia et al., 2016). Young children are especially vulnerable to infectious diseases and injuries because of their age-appropriate behavior and abilities, their immature immune systems, and their lack of understanding of risk. Maintaining safe and healthful environments and practices involves removal of hazards and provision of policies and procedures, as well as compliance with quality standards by everyone in the group.

Numerous studies have shown the effectiveness of child care health consultation. This study focused on I/T care. The immediate intervention group showed significant improvement in policy development for safe sleep and child abuse and in education about safe sleep practices, preventing child abuse, and medication administration training. Some improvement in diaper changing and hand hygiene procedures occurred. The delayed intervention (contract) group showed significant improvement in safe sleep procedures, policies

and education, medication administration procedure, diaper changing procedures, and care plans for children with special needs with appropriate information and signed by a health care provider.

The data collected by ITQIP show that many children with special needs lacked appropriate care plans. After finding little improvement in the immediate intervention group for centers having care plans with needed elements, ITQIP chose this topic as the focus of an MCHB-required continuous quality improvement initiative. ITQIP provided an audioconference for the CCHCs and gave them resources for teaching what should be in a care plan. CCHCs reported that they were most successful at helping the centers have complete, useful care plans for children with disease-specific conditions.

The areas chosen to target varied from center to center. Immunization was chosen by only one center. At the time of the study, neither regulation inspectors nor quality rating assessors were checking whether the center had documentation that the enrolled children were up to date with their vaccines. With little incentive or sanctions, documentation of up-to-date immunization status was poor.

Improvements occurred in some practices specified in selected *CFOC3* standards. Many of the directors said they appreciated the help they received from the CCHCs that ITQIP linked with their centers. The director of one center, part of a corporation with centers in 12 states, advocated for improving sleep policies for all the centers in her company. This advocacy could lead to widespread improvement.

The centers that participated in this project were STAR 2 and STAR 3 programs that responded to an invitation to participate in ITQIP to improve. They were willing to contribute a modest copayment to work with a CCHC and wanted to raise their STAR rating and consequent higher payments for subsidized enrollees.

The immediate intervention group showed significant improvement in policy development for safe sleep and child abuse and in education about safe sleep practices, preventing child abuse, and medication administration training.

Many of the directors said they appreciated the help they received from the CCHCs that ITQIP linked with their centers.

This selection bias is likely to have influenced the observed improvements.

A limitation of the study is the small sample size due to limited funding for the project. Also, although the study assessed practices for 13 *CFOC3* standards (AAP et al., 2011), the centers addressed only three topic areas. Little improvement was seen in topics that were not chosen or chosen less frequently. Change in leadership at the centers with varying levels of interest in working on the action plans made improvement difficult.

Another limitation of the study is the variability in child care operation from one facility to another and from year to year. Evaluators were unlikely to have been evaluating the same children from pretest through Posttest 2. Different teachers/caregivers and children may occupy designated rooms in a facility. ITQIP did not require that the CCHCs spend a specific amount of time with their centers. The time and type of service provided by CCHCs varied widely. Although CCHCs reported the total time and types of services they provided, they were not asked to report the time spent in each type of service (onsite visits, phone calls, or e-mails).

CCHCs support health and safety practices and environments that prevent harm and promote health and development of children, as well as overall wellbeing for families and early education staff. Currently, only 17 U.S. states have a statutory requirement for early childhood education programs to have child care health consultation (Honigfeld, Pascoe, Macary, & Crowley, 2017). Of these, two states require CCHC involvement only if the facility cares for sick children (Honigfeld et al., 2017).

None of the centers in this project continued their relationship with their CCHC after the year of subsidized linkage. Some directors stated that although they found the CCHC very helpful and informative, the cost of the CCHC was prohibitive. Some said they would continue the CCHC on a fee basis if they could budget for it in the future. Other studies have shown that linkage of centers with CCHC improves health and safety compliance. ITQIP showed this is true for I/T programs, too.

REFERENCES

- Alkon, A., & Bernzweig, J. (2008). *Child care health linkages project evaluation summary*. California Childcare Health Program, UCSF School of Nursing, September 2001 – June 2004. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1525-1446.2008.00689.x/full>
- Alkon, A., Bernzweig, J., Kim, T., Mackie, J., Wolff, M., & Elman, J. (2008). Child care health consultation programs in California: Models, services, and facilitators. *Public Health Nursing*, 25, 126-139.
- Alkon, A., Bernzweig, J., Kim, T., Wolff, M., & Mackie, J. (2009). Child care health consultation improves health and safety policies and practices. *Academic Pediatrics*, 9, 366-370.
- Alkon, A., Crowley, A., Benjamin Neelon, S., Hill, S., Pan, Y., Nguyen, V., ..., Kotch, J. B. (2014). Nutrition and physical activity randomized control trial in child care centers improves knowledge, policies, and children's body mass index. *BMC Public Health*, 14, 215.
- Alkon, A., Nouredini, S., Swartz, A., Sutherland, A. M., Stephens, M., Davidson, N., & Rose, R. (2016). Intergrated Pest Management Intervention in child care centers improves knowledge, pest control, and practices. *Journal of Pediatric Health Care*, 30, e27-e41.
- Alkon, A., Sokal-Gutierrez, K., & Wolff, M. (2002). Child care health consultation improves health knowledge and compliance. *Pediatric Nursing*, 28, 61-65.
- American Academy of Pediatrics, American Public Health Association, National Resource Center for Health and Safety in Child Care and Early Education. (2011). *Caring for our children: National health and safety performance standards; Guidelines for early care and education programs* (3rd ed.). Elk Grove Village, IL: American Academy of Pediatrics; Washington, DC: American Public Health Association.
- American Academy of Pediatrics, American Public Health Association, National Resource Center for Health and Safety in Child Care and Early Education. 2013. *Stepping stones to caring for our children: National health and safety performance standards; Guidelines for early care and education programs* (3rd ed.). Elk Grove Village, IL: American Academy of Pediatrics; Washington DC: American Public Health Association.
- Banghart, P., & Kraeder, J. L. (2012). *What can CCDF learn from the research on children's health and safety in child care? Urban Institute Brief #03*. Washington, DC: Urban Institute. Retrieved from <http://www.urban.org/publications/412579.html>
- Carabin, H., Gyorkos, T., Soto, J., Joseph, L., Payment, P., & Collet, J. P. (1999). Effectiveness of a training program in reducing infections in toddlers attending day care centers. *Epidemiology*, 10, 219-227.
- Crowley, A. (2006) *Improving child and family health through child care health consultation. Executive summary*. Beatrice Renfield-Yale School of Nursing Clinical Research Initiatives Fund. Yale University School of Nursing. New Haven, CT: Author.
- Garcia, J. L., Heckman, J. J., Leaf, D. E., & Padros, M. J. (2016). *The life-cycle benefits of an influential early childhood program. Working paper 2016-035*. Chicago, IL: University of Chicago. Retrieved from <https://heckmanequation.org/resource/research-summary-lifecycle-benefits-influential-early-childhood-program/>
- Harms, T., Cryer, D., & Clifford, R. (2006). *Infant/toddler environmental rating scale (revised ed.)*. New York: Teachers College Press.
- Honigfeld, L., Pascoe, T., Macary, S., & Crowley, A. (2017). Promoting children's health in early care and education settings by supporting health consultation. *Impact*. Farmington, CT: Child Health and Development Institute of Connecticut.
- Isbell, P., Kotch, J., Savage, E., Gunn, E., Lu, L., & Weber, D. (2013). Improvement of child care programs' policies, practices, and children's access to health care linked to child care health consultation. *NHSA Dialog: A Research to Practice Journal*, 16, 34-52.
- Kotch, J. B. (2002). *The Quality Enhancement Project for Infants and Toddlers: Executive summary*. Chapel Hill, NC: University of North Carolina at Chapel Hill.
- Kotch, J. B., Isbell, P., Weber, D. J., Nguyen, V., Savage, E., Gunn, E., ..., Allen, J. (2007). Hand-washing and diapering equipment reduces disease among children in out-of-home child care centers. *Pediatrics*, 120, e29-e36.
- Laughlin, L. (2013). Who's minding the kids? Child care arrangements: Spring 2013. Current Population Reports, P70-135. Washington, DC: U.S. Census Bureau. Retrieved from <https://www.census.gov/prod/2013pubs/p70-135.pdf>
- Moon, R., & Oden, R. (2005). Back to sleep: Can we influence child care providers. *Pediatrics*, 112, 878-882.

- Organizational Research Services and Geo Education & Research. (2007). Child care health consultation: Evidence based effectiveness. Data from Healthy Child Care Washington evaluation report 2003–2007. Seattle, WA: Washington State Department of Health. Retrieved from https://www.napnap.org/sites/default/files/userfiles/membership/CCS_SIG_Evidence_%20Based_%20CCHP.pdf
- Pacific Research and Evaluation. (2008). *Improving the health and safety of children in Oregon's child care: Implementation and outcomes of Oregon Child Care Health Consultation Program*. Portland, OR: Department of Human Services, Public Health Division, Office of Family Health.
- Pacific Research and Evaluation. (2007). *Evaluation of the Child Care Health Consultation Demonstration Program: Phase IV Final Report 2007*. Portland, OR: Department of Human Services, Office of Family Health.
- Pennsylvania Department of Human Services. (2008). 55 Pa Code, Chapter 3270, Child Care Regulations; 3270.131(a)(5). Retrieved from <http://www.dhs.pa.gov/provider/earlylearning/childcareregulations/>
- Pennsylvania Key Program Quality Assessment Team. (2016). [2015–2016 Environmental rating scale reports: ECERS-R, ECERS-3, ITERS-R]. Unpublished raw data.
- Pennsylvania Office of Child Development and Early Learning. (2010). *Demonstrating quality: Pennsylvania Keystone STARS 2010 Program report*. Harrisburg, PA: Author. Retrieved from <http://www.ocdelresearch.org/Reports/Forms/AllItems.aspx?RootFolder=%2FReports%2FKeystone%20STARS&FolderCTID=0x01200092EA27E29EEE3E4AAE2D4C5508AC9E5A&View={5EEC6855-F8A8-486E-B6E0-FE6B9FDEBE2E}>
- Ramler, M., Nakatsukasa-Ono, W., Loe, C., & Harris, K. (2006). *The influence of child care health consultants in promoting children's health and well-being: A report on selected resources*. Rockville, MD: Maternal Child Health Bureau.
- Roberts, L., Smith, W., Jorm, L., Patel, M., Douglas, R. M., & McGilchrist, C. (2000a). Effect of infection control measures on the frequency of upper respiratory infection in child care: A randomized, controlled study. *Pediatrics*, 105, 738-742.
- Roberts, L., Smith, W., Jorm, L., Patel, M., Douglas, R. M., & McGilchrist, C. (2000b). Effect of infection control measures on the frequency of diarrheal episodes in child care: A randomized, controlled study. *Pediatrics*, 105, 743-746.
- Weinburg, S. (2017). WellCareTracker™ [Computer software]. Retrieved from www.wellcaretracker.org
- Zaslow, M., Tout, K., & Halle, T. (2012). *On-site approaches to quality improvement in quality rating and improvement systems: Building on the research on coaching. Research-to-Policy, Research-to-Practice Brief OPRE 2012-40*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Comparison of State Online Mandated Reporter Trainings

*Danielle Somerville, BS; Sarah Dore, MPA;
Kathryn L. Humphreys, PhD, EdM;
Nicole Verdiglione, MA; Shouhau Zhou, PhD;
Robert Hamm, PhD; Rick Fiene, PhD; Karl Kapp, EdD;
Benjamin H. Levi, MD PhD FAACP*

Key words: Mandated reporter training; Online learning; Child abuse, Child protection; Evidence-based training

Abstract

This article presents the results of a comparative analysis of online mandated reporter trainings regarding child abuse. Programs from 47 U.S. states and the District of Columbia were reviewed and their content and features compared with iLookOut for Child Abuse's Core Training. Significant variation was identified in terms of the scope, content, didactic approach, delivery method, and outcome measures across different trainings. These findings raise concern that while all children need protection from abuse, not all mandated reporters are receiving comparable preparation to fulfill this important moral and legal responsibility.

Introduction

As individuals required by law to report suspected *child abuse* (a term used here to denote all forms of child maltreatment, including neglect), mandated reporters play an important role in protecting victimized and at-risk children (Child Welfare Information Gateway, 2019). Though most states require mandated reporters to complete state-approved training on how to recognize and report suspected child abuse, there are currently no standards for such training in terms of content or delivery (Kenny et al., 2016). The lack of national consensus for child abuse training along with the potential for major differences in training across jurisdictions raise concern that not all mandated reporters are receiving comparable

preparation and that some are being inadequately prepared (Kenny, 2015). Many states use online training for instruction in mandated reporting, but little is known about how these trainings vary in terms of content and delivery. This article reports on the findings of such a comparison and discusses some of the implications of the variability that was found.

Research has shown that online training is effective for educating adults and that it has particular advantages over in-person training (John et al., 2020). Specifically, online training is accessible (in terms of both timing and content), enables standardized evaluations of learning and satisfaction, facilitates storage and analysis of data, and can readily promote continuous learning (Kenny et al., 2016; Scott et al., 2016; Shendell et al., 2016).

Mandated Reporter Training

Particularly when in-person training is not feasible, well-designed evidence-based online training can ensure that quality education is available to a multitude of people.

In response to the need for evidence-based online mandated reporter training (Ayling et al., 2019), a research team at the Penn State College of Medicine developed the iLookOut for Child Abuse (iLookOut) Core Training—a free online, interactive learning program designed to help mandated reporters better recognize, report, and respond to suspected child abuse (www.ilookoutproject.org). iLookOut also delivers an Advanced Training module. Although the iLookOut training was initially developed for early childhood professionals, it includes general information to meet the needs of all mandated reporters (e.g., in Pennsylvania) and can be adapted for use in any state. iLookOut's Core Training uses an experiential learning framework, a video-based storyline, and gamification to engage users (Levi et al., 2019). In addition to a standard registration page, the iLookOut Core Training includes validated pre/post measures that evaluate (1) knowledge and attitudes about child abuse and its reporting, (2) individuals' sense of preparedness (to identify and report suspected abuse), and (3) learners' satisfaction with the iLookOut Core Training. In-depth descriptions have been published on iLookOut's content and structure, practical and conceptual considerations in creating iLookOut, and its strategy for ensuring that its educational material is fully integrated (Kapp et al., 2020; Levi et al., 2019; Levi et al., in press). Also published are data from a randomized controlled trial (n=741) and a real-world study (n=11,065) demonstrating that the iLookOut Core Training significantly improves knowledge and changes attitudes regarding child abuse and its reporting compared with a standard mandated reporter training (Humphreys et al., n.d.; Mathews et al., 2017; Yang et al., 2020). While the iLookOut learning program's effects are well documented, there is little research on the effectiveness of other online mandated report trainings. In light of these findings, we sought to compare iLookOut's Core Training with

other existing online mandated reporting trainings in all 50 U.S. states.

Methods

A primary reviewer from the iLookOut research team performed a comparative analysis of 48 online mandated reporter trainings, including the iLookOut Core Training, and the findings were then reviewed and confirmed. Specifically, the primary reviewer registered for and completed (in full) each and every online training examined, and the reviewer then binary coded (present or absent) each of the 40 characteristics for every training. The process for generating the list of characteristics is described in the next section. Each training was further assessed for overall Level of Engagement, based on the presence or absence of several interactive features (see Table 4) as well as the scope of information present in each training. The initial coding process was completed by the primary reviewer, cross-checked for accuracy by two additional reviewers (no disagreements were identified), and discussed and confirmed according to the findings by a larger multi-disciplinary team.

Identification of Training

The initial step to identify online mandated reporter (MR) training for each state involved querying the Child Welfare Information Gateway (2019), a service of the U.S. Department of Health and Human Services that provides online resources to professionals in child welfare and related fields. Their webpage, State Training Resources for Mandatory Reporters of Child Abuse and Neglect, provides a state-by-state list of online MR trainings and other resources (e.g., toolkits, guidelines, protocols) for mandated reporters of child abuse. For the 35 states for which this listing identified a specific state-sponsored, publicly available MR online training, that training was used for the present comparison.

For states where no such program was identified, a subsequent search was performed for trainings sponsored by non-profit agencies whose primary focus was child protection/child abuse prevention, such as CARE House, Michigan (CARE House of Oakland County, 2020). If this search did not identify an online MR training, a new search was conducted for MR trainings offered by more broad-based organizations, such as SafeSchools Training, Ohio (SafeSchools, 2020), for whom child protection was not the primary focus. For states in which there were multiple state-*approved* MR trainings, but no state-*sponsored* MR trainings, the state-approved training that was both most accessible (e.g., free, user-friendly) and most extensive (in terms of content and delivery/presentation) was selected for inclusion and comparison. The various state trainings and their classifications are shown in Tables 1 and 2.

Inclusion/Exclusion Criteria

Only MR trainings that were publicly available online (either open access or with registration) were examined. With the exception of Kansas, Nebraska, and Oklahoma (which charged \$5, \$15, and \$15, respectively), none of the trainings required a fee for access or to obtain a certificate of completion. All sites were accessed between January 3, 2020, and October 26, 2020. To enable maximal comparability, only English-version trainings were evaluated; however, Table 2 identifies states that also provided MR training in multiple languages.

Subject Matter for Comparison

According to expert recommendations (Damashek et al., 2011; Chen et al., 2013), effective MR training should (1) imbue *knowledge* about the various types of child abuse, risk factors, and the long-term epidemiology and impact of child abuse and also (2) cultivate *skills* for recognizing both physical signs of abuse and behavioral indicators of child abuse (for both children and perpetrators). Based on published

recommendations, and using iLookOut as the reference training, an iteratively constructed list of 40 characteristics was created to account for the kinds of content and functionality present in each training. The list was developed by the primary reviewer in collaboration with a multi-disciplinary team whose research focuses on child protection. Content, which accounted for 21 of those 40 characteristics, included both fact-based information (e.g., types of abuse, risk factors for abuse, legal responsibilities of mandated reporters, steps for making a report) as well as education about processing (e.g., how to ask better questions, respond to disclosures by children, interpret “reasonable suspicion”). Matters involving Delivery/Functionality accounted for the remaining 19 characteristics and included the presence of a pre-and/or post-test, handouts, videos, voice narration, links to resources, as well as elements that promoted engagement (e.g., user-friendly format, interactive games, stories). A full list and description of these 40 characteristics can be found in Table 3.

MR trainings were then categorized as Limited, Basic, Moderate, or Advanced based on their level of engagement, as per the inclusion criteria shown in Table 4. Because there are no published consensus criteria for grading MR trainings, these three categories were intended to broadly categorize the different tiers of training as they currently stand. It is our hope that the present examination encourages others in the field to develop evidence-based, outcomes-driven criteria for a quality rating system of MR trainings.

Results

Format

In addition to the iLookOut Core Training, MR trainings were identified and examined for all U.S. states plus Washington, D.C., with the exception of Mississippi, Rhode Island, and Wyoming, for which no online MR trainings were identified. A list of all 49 trainings examined and the agencies and organizations that sponsored each training are

Mandated Reporter Training

included in Table 2. The amount of time it took to complete each MR training ranged from 30 minutes to 3 hours. While some of the trainings allowed users to click through modules at their own pace, others required users to remain in a given module

for a fixed amount of time. Most provided some form of overview to orient the user (n=32), and most included video-based content (n=32). However, over one third of MR trainings (n=17) presented information using only slides or text.

Table 1. Providers of Online Mandated Reporter Trainings.

Type of training	Number of states	Which states
State-sponsored mandated reporter trainings (*provided by state university)	35	AL, AK, AZ, AR, CA, CO, CT, DC, DE, FL, HI, IL, IN, IA, KS, LA, ME, MD, MN, NV, NJ, NM*, NY, ND, OK, OR, PA, SC, SD, TN, TX, VT, VA, WA, WI
Non-profit, primary focus on child abuse	11	ID, KY, MA, MI, MO, MT, NC, NE, NH, UT, WV,
Non-profit, broader focus on child well-being	2	GA, OH
No online training found	3	MS, RI, WY

Table 2. State Mandated Reporting Training Sources.

State	Agency
Alabama	Alabama Dept. of Human Resources https://aldhr.remote-learner.net/ Last accessed: July 2020
Alaska	Alaska Department of Health and Human Services http://training.dhss.alaska.gov/mandatoryreporter/training/multiscreen.html Last accessed: July 2020
Arizona	Arizona Child Abuse Info Center – Children’s Justice Program https://childhelpinfocenter.org/ Last accessed: July 2020
Arkansas	Arkansas Commission on Child Abuse, Rape and Domestic Violence https://ar.mandatedreporter.org/UserAuth/Login!loginPage.action Last accessed: July 2020
California	California Department of Social Services https://mandatedreporterca.com/ (also available in Spanish) Last accessed: August 2020
Colorado	Colorado Department of Human Services https://coloradocwts.com/mandated-reporter-training Last accessed: July 2020
Connecticut	Connecticut Department of Children and Families https://www.proprofs.com/training/course/?title=connecticut-mandated-reporter-training-for-community-providers-jan-2020-version-3_5e260a8c470e8 Last accessed: July 2020
Delaware	Delaware Office of the Child Advocate https://ocade.server.tracorp.com/novusii/application/login/

Mandated Reporter Training

State	Agency
District of Columbia	District of Columbia Children and Family Services Agency https://dc.mandatedreporter.org/pages/Welcome.action Last accessed: July 2020
Florida	Florida Department of Children and Families https://fl-dcf.org/RCAAN/_media/RCAAN/index.html#SPLASH Last accessed: July 2020
Georgia	Georgia Division of Family and Children's Services https://www.prosolutionstraining.com/store/product/?tProductVersion_id=1093 (also available in Spanish) Last accessed: August 2020
Hawaii	Department of Human Services- Social Services https://humanservices.hawaii.gov/ssd/home/child-welfare-services/ Last accessed: July 2020
Idaho	IdahoStars https://idahostars.org/portals/61/Docs/Providers/ApprovedTrain/ICCP/ICCP_ReportAbuse_Inst_2018.pdf Last accessed: July 2020
Illinois	Illinois Department of Children and Families https://mr.dcfstraining.org/UserAuth/Login!loginPage.action Last accessed: July 2020
Indiana	Indiana Department of Child Services https://reportchildabuse.dcs.in.gov/ Last accessed: August 2020
Iowa	Iowa State University Extension and Outreach https://training.hs.iastate.edu/course/view.php?id=731#section-2 Last accessed: July 2020
Kansas	Kansas Child Care Training Opportunities https://kccto.org/product/strengthening-families-through-positive-connections/ Last accessed: August 2020
Kentucky	Kosair Charities® https://education.ky.gov/teachers/Documents/CANtraining_FaceIt.pdf Last accessed: April 2020
Louisiana	Louisiana Department of Children and Families http://www.dcf.louisiana.gov/index.cfm?md=newsroom&tmp=detail&articleID=575#undefined Last accessed: April 2020
Maine	Maine Office of Child and Family Services https://www.maine.gov/dhhs/ocfs/mandated-reporters.shtml Last accessed: April 2020
Maryland	Maryland's Resource for Mandated Reporters https://training.reportabusemd.com/ Last accessed: August 2020
Massachusetts	Middlesex Children's Advocacy Center https://51a.middlesexcac.org/ Last accessed: April 2020
Michigan	CARE House of Oakland County https://mandatedreportertraining.carehouse.org/welcome/?profession=1 Last accessed: July 2020
Minnesota	Minnesota Department of Human Services https://mn.gov/dhs/people-we-serve/children-and-families/services/child-protection/programs-services/mandated-reporting-training-overview.jsp Last accessed: July 2020
Missouri	Missouri Kids First https://protectmokids.com/ Last accessed: July 2020
Montana	Child Care Resources, Inc. https://www.childcaretraining.org/mod/page/view.php?id=4007 Last accessed: July 2020
Nebraska	Project Harmony https://projectharmony.learnupon.com/store?utf8=%E2%9C%93&ss=1&ct=93426&commit=Filteraining.org/?pageid=84 Last accessed: September 2020

Mandated Reporter Training

State	Agency
Nevada	The Nevada Registry http://www.nevadaregisNtry.org/ Last accessed: April 2020
New Hampshire	Know and Tell https://knowandtell.org/educate/ Last accessed: April 2020
New Jersey	New Jersey Department of Education https://www.state.nj.us/education/students/safety/socservices/abuse/training/ Last accessed: April 2020
New Mexico	New Mexico State University https://swrtc.nmsu.edu/educators/
New York	New York State Office of Children and Family Services http://www.nysmandatedreporter.org/ Last accessed: April 2020
North Carolina	Prevent Child Abuse North Carolina https://preventchildabusenc-lms.org/ (also available in Spanish) Last accessed: April 2020
North Dakota	North Dakota Department of Human Services http://www.pcand.org/NDDHS/mandatedreportertraining/index.html Last accessed: April 2020
Ohio	SafeSchools Training https://www.safeschools.com/courses/child-abuse-mandatory-reporting-ohio/ Last accessed: April 2020
Oklahoma	The University of Oklahoma Health Sciences Center https://www.ouhsc.edu/okcantraining/Online-Training Last accessed: July 2020
Oregon	Oregon Department of Human Services https://www.oregon.gov/DHS/ABUSE/Pages/mr_employees.aspx Last accessed: April 2020
Pennsylvania	Pennsylvania KeepKidsSafe https://www.reportabusepa.pitt.edu/webapps/portal/execute/tabs/tabAction?tab_group_id=911 Last accessed: April 2020
South Carolina	University of South Carolina School of Law https://dss.sc.gov/child-well-being/mandated-reporters/ Last accessed: April 2020
South Dakota	South Dakota Department of Social Services https://apps.sd.gov/SS60ReporterVideoTraining/Introduction.aspx Last accessed: September 2020
Tennessee	Tennessee State Government https://www.tn.gov/dcs/program-areas/training/tpd/cw-resources/cwr/mandated-reporter-training.html Last accessed: October 2020
Texas	Texas Dept. of Family and Protective Services http://www.dfps.state.tx.us/training/reporting/ (also available in Spanish) Last accessed: July 2020
Utah	Prevent Child Abuse Utah https://pcautah.org/ Last accessed: July 2020
Vermont	KidsSafe Collaborative, Vermont Agency of Human Services https://goto.webcasts.com/starthere.jsp?ei=1087433 Last accessed: April 2020
Virginia	Virginia Department of Social Services https://www.dss.virginia.gov/abuse/mr.cgi
Washington	Washington State Department of Children, Youth and Families https://www.dcyf.wa.gov/safety/mandated-reporter Last accessed: April 2020
West Virginia	Prevent Child Abuse West Virginia https://teamwv.org/prevent-child-abuse-wv-landing/mandated-reporter-training-information/ Last accessed: April 2020
Wisconsin	Wisconsin Child Welfare Professional Development System https://media.wcwpds.wisc.edu/mandatedreporter/ Last accessed: April 2020

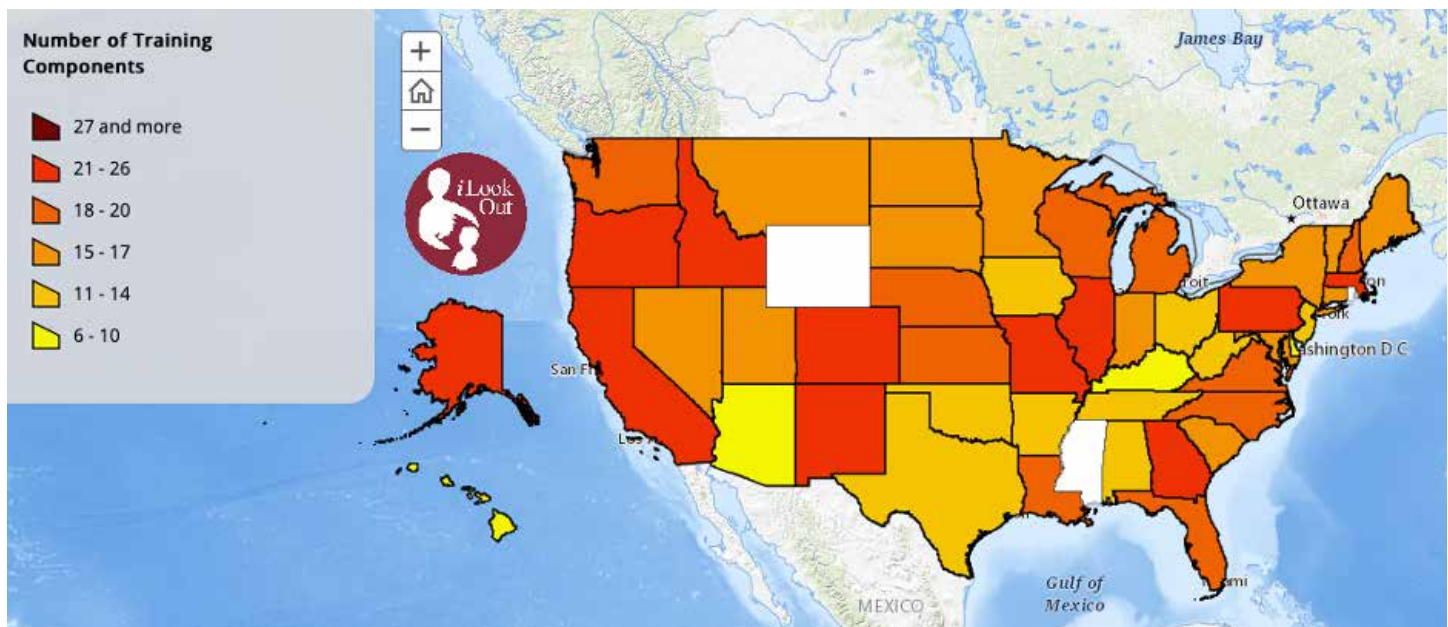
Testing

Knowledge checks (i.e., fact-based questions) were present at various junctures in 30 of the MR trainings, while only nine MR trainings included *skill-testing activities* (e.g., recognizing abuse, identifying risk factors for abuse). Though 30 of the MR trainings included a *post-test* to evaluate user knowledge, only 12 of these contained a *pre-test* such that they could measure pre-/post-test changes in knowledge. *Real-world scenarios* in the form of short stories and multiple-choice questions (Errington, 2008) were present in 24 MR trainings, and 32 trainings included some element of voice narration, but only seven MR trainings provided any form of extended *scenario-based storyline*.

Content

MR trainings also varied considerably in terms of specific content. The vast majority (n=42) provided detailed information about the process for *making a report* of child abuse, but only 10 MR trainings provided explanations about interpreting the statutory threshold for when mandated reporting is required (colloquially referred to as *reasonable suspicion*) (Levi & Loeben, 2004). Relatively few MR trainings included information about *domestic violence* or *animal abuse* being risk factors for child abuse (n=13), how to determine when *reasonable suspicion* is present (n=8), or what kinds of questions are *better* (n=17) or *worse* (n=14) to ask when responding to a child's disclosure of abuse.

Figure 1. Number of Training Components by State.



(States colored white do not have an online MR training.)

Mandated Reporter Training

Table 3. Training Components and Their Presence in State MR Trainings.

Component	Definition	States that have this component	States
Training features			
Pre-test	Pre-training fact-based questionnaire that tested knowledge about child abuse and its reporting	14 (29%)	AZ, AR, DC, GA, IL, IA, KS, LA, MO, MT, NH, OK, UT, iLookOut
Post-test	Post-training fact-based questionnaire that tested knowledge about child abuse and its reporting	30 (61%)	AL, AK, AZ, AR, CO, DC, GA, ID, IL, IN, KS, LA, ME, MD, MI, MO, NE, NV, NH, NC, OH, OK, OR, PA, SC, TN, TX, UT, VT, iLookOut
Training overview	Introduction that explains the purpose, content, and format of the training	32 (65%)	AK, CA, CO, DE, DC, FL, GA, ID, IL, IN, KS, LA, MA, MI, MN, MO, MT, NE, NH, NJ, NY, OH, OR, OK, PA, SC, TX, VT, VA, WV, WI, iLookOut
Individualized learning path	Individualized pathway that learners choose to guide their training content	2 (4%)	CO, ND
PDF handouts	Informational handouts that can be downloaded	10 (20%)	FL, ID, IA, MO, NJ, OK, OR, WA, WI, iLookOut
Reporting worksheet	Form that can be downloaded and used to take notes in preparation for making a report	4 (8%)	FL, OR, WI, iLookOut
Questions/knowledge checks throughout training	Short quizzes that follow each lesson	30 (61%)	AL, AR, CA, CO, CT, DC, GA, IL, IN, IA, KS, ME, MA, MI, MO, MT, NE, NV, NJ, NM, NC, ND, OR, PA, SC, TX, UT, VT, VA, iLookOut
Skill testing activities	Interactive games or activities (i.e., matching activity, crossword puzzle) that test knowledge of various topics	9 (18%)	CO, GA, MO, NM, NY, OR, SC, WI, iLookOut

Mandated Reporter Training

Component	Definition	States that have this component	States
Feedback/explanation after questions	Explanations of correct/incorrect answers following knowledge checks	18 (37%)	AL, CA, CO, CT, DC, ID, IN, ME, MA, MO, NE, NM, OH, OR, PA, UT, VT, iLookOut
Real-world scenarios	Real-world applications that provide context for training content	24 (49%)	AL, AK, AR, CO, DE, GA, IL, IN, LA, MI, MT, NE, NH, NM, NC, ND, OR, PA, SC, SD, UT, VA, WI, iLookOut
Scenario-based storyline	A storyline used to engages learners in the training	7 (14%)	AL, AK, MI, NM, ND, WI, iLookOut
Voice narration	A spoken commentary accompanies text throughout the training	32 (65%)	AL, AK, CA, CO, CT, DC, FL ID, KY, LA, ME, MD, MA, MI, MN, MO, NE, NV, NY, NC, ND, OK, OR, SD, TN, TX, UT, VT, VA, WA, WI, iLookOut
Videos	Informational videos included throughout the training	32 (65%)	AL, AK, AZ, CA, CO, DC, HI, ID, KY, MD, MA, MI, MN, MO, MT, NE, NV, NH, NM, NY, ND, OH, OK, OR, SC, SD, TN, TX, VT, WA, WI, iLookOut
Videos with real actors	Video scenarios that are acted out by live actors	7 (14%)	AK, KS, NH, NM, TX, VT iLookOut
Audio clips	Short voice-clips to narrate slides and/or modules	12 (24%)	CO, CT, IL, IN, IA, ME, MO, NJ, NM, NC, WI, iLookOut
Links to online resources	Links to additional resources such as state laws, definitions, or further information about the topics covered in training	21 (42%)	AK, CA, CO, CT, DC, FL, GA, ID, IL, IN, KS, MD, MI, NH, NM, OR, PA, SC, SD, TX, iLookOut
Resources for ongoing training (ie, micro-learning)	Resources that promote on-going learning beyond the initial training (e.g., follow-up micro-learning activities)(18).	1 (2%)	iLookOut

Mandated Reporter Training

Component	Definition	States that have this component	States
Evaluation of mandated reporter training	A survey given at the end of the training to gauge the user's overall experience	15 (31%)	DC, GA, ID, IL, KS, LA, NE, NH, NC, PA, SC, SD, UT, VT, iLookOut
Discussion forum	An online forum that allows for ongoing discussion between learners	1 (2%)	KS
Mandated reporter content	Content		
Types of abuse	Definition of the main types of child abuse/neglect, as state definitions	48 (98%)	AL, AK, AZ, AR, CA, CO, CT, DE, DC, FL, GA, ID, IL, IN, IA, KS, KY, LA, ME, MD, MA, MI, MN, MO, MT, NE, NV, NH, NJ, NM, NY, NC, ND, OH, OK, OR, PA, SC, SD, TN, TX, UT, VT, VA, WA, WV, WI, iLookOut
Common signs and symptoms of abuse	Common signs and symptoms of each of the main types of child abuse/neglect	44 (90%)	AL, AK, AR, CA, CO, CT, DE, DC, FL, GA, HI, ID, IL, IN, IA, KS, LA, ME, MD, MA, MI, MN, MO, MT, NE, NV, NH, NJ, NM, NY, NC, ND, OH, OR, PA, SC, SD, TN, VT, VA, WA, WV, WI, iLookOut
Risk factors	Factors that put certain children/families at greater risk for experiencing maltreatment	20 (41%)	AR, CA, CO, CT, DC, GA, ID, IL, KY, LA, ME, MA, MO, NE, OR, PA, UT, VT, VA, iLookOut
Red flags/indicators for abuse	Physical/behavioral indicators that should raise concern about child abuse	41 (84%)	AL, AK, CA, CO, CT, DC, FL, GA, ID, IL, IN, IA, KS, KY, LA, ME, MD, MA, MI, MN, MO, MT, NE, NV, NH, NJ, NM, NY, NC, ND, OH, OR, PA, SC, SD, UT, VT, VA, WA, WV, iLookOut

Mandated Reporter Training

Component	Definition	States that have this component	States
Things that should NOT raise concern	Physical/behavioral signs that should NOT raise concern about child abuse (i.e., Mongolian spots, normal locations where toddlers bruise)	17 (35%)	AR, CA, CO, GA, IL, KS, ME, MN, MT, NM, NC, ND, VT, VA, WA, WI, iLookOut
Parent/caregiver behavior that should raise concern	Behavior seen in child care providers that should raise concern about possible abuse	24 (49%)	AK, CA, DE, DC, FL, ID, KS, KY, LA, MI, MO, NE, NV, NH, NY, NC, ND, OR, PA, SC, UT, VT, WA, iLookOut
Prohibited child care provider behavior	Behavior seen in parents or caretakers that should raise concern about possible abuse	1 (2%)	ID
Epidemiology of child abuse	Statistics that describe the scope of child maltreatment on a national or state level	25 (51%)	AK, CA, DE, DC, GA, IN, IA, KS, LA, ME, MD, MA, MO, NE, NV, NH, NM, OH, OR, PA, TN, UT, VA, WV, iLookOut
Long-term impact of child abuse	Long-term physical, psychological or behavioral consequences of child maltreatment, as well as costs to society as a whole	23 (47%)	AL, AK, CA, DE, GA, ID, KS, LA, ME, MD, MA, MI, MN, MO, NV, NH, ND, OH, OR, PA, UT, WV, iLookOut
Domestic violence/animal abuse	Domestic violence and/or animal abuse presented as risk factors for abuse	13 (27%)	CA, CT, DC, IL, ME, MN, MT, NH, NM, NC, VT, WA, iLookOut
Explication of reasonable suspicion mean	Explanation of what "reasonable suspicion" means (with regard to making a report of suspected abuse)	10 (20%)	AZ, CT, FL, MA, MO, NM, NY, OK, PA, iLookOut
Determining whether reasonable suspicion is present	Examples are given of how to determine whether or not a situation rises to the level of reasonable suspicion	8 (16%)	FL, ID, MT, NY, OK, PA, VT, iLookOut

Mandated Reporter Training

Component	Definition	States that have this component	States
Information gathering—good questions	Examples of better questions to ask when responding to a disclosure of child maltreatment from a child	17 (35%)	AK, AZ, FL, GA, IL, KY, ME, MA, MO, NE, NH, UT, VA, WA, WV, WI, iLookOut
Information gathering—bad questions	Examples of worse questions to ask when responding to a disclosure of child maltreatment from a child	14 (29%)	AK, GA, ID, IL, MA, MO, NE, NH, TN, VA, WA, WV, WI, iLookOut
How to respond to disclosures by children	How to talk to a child who has disclosed that they have been abused	32 (65%)	AK, AZ, AR, CA, CO, DE, DC, FL, GA, ID, IL, KY, LA, MD, MA, MO, MT, NE, NV, NH, NM, NC, OH, SC, SD, TN, UT, VA, WA, WV, WI, iLookOut
Legal responsibilities of mandated reporters	State laws that designate which professionals are required to report cases of suspected child abuse/neglect	44 (90%)	AK, AZ, AR, CA, CO, CT, DC, FL, GA, HI, ID, IL, IN, IA, KS, LA, ME, MD, MA, MI, MN, MO, MT, NE, NV, NH, NM, NY, NC, OH, OK, OR, PA, SC, SD, TN, TX, UT, VT, VA, WA, WV, WI, iLookOut
Consequences for failing to report	Penalties for mandated reporters who fail to report cases of suspected child abuse/neglect	37 (76%)	AZ, AR, CO, CT, DC, FL, GA, HI, ID, IL, IN, IA, KS, LA, ME, MD, MA, MI, MN, MO, MT, NE, NV, NJ, NY, NC, OH, OK, OR, PA, SC, SD, TN, VT, VA, WA, iLookOut
Legal protection for good faith reports	Explanation that mandated reporters are protected from liability if a report is made in good faith, regardless of the outcome of the report	41 (84%)	AL, AK, AZ, AR, CA, CO, CT, FL, GA, HI, ID, IL, IN, IA, LA, ME, MD, MA, MI, MN, MO, MT, NE, NV, NJ, NM, NY, NC, ND, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, WI, iLookOut

Mandated Reporter Training

Component	Definition	States that have this component	States
Preparing to make a report	Specific information that should be gathered before making a report of suspected child abuse	43 (88%)	AL, AK, CA, CO, CT, DE, DC, FL, GA, HI, ID, IN, IA, KS, LA, ME, MD, MA, MI, MN, MO, MT, NV, NH, NJ, NM, NY, NC, ND, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, WA, WV, WI, iLookOut
Mechanics of making a report	Specific steps involved in making a report of suspected abuse to child protective services	42 (88%)	AL, AK, AR, CA, CO, CT, DE, DC, FL, GA, ID, IL, IN, IA, KS, LA, MD, MA, MI, MN, MO, MT, NE, NV, NH, NJ, NM, NY, NC, ND, OK, OR, PA, SC, SD, TN, TX, VA, WA, WV, WI, iLookOut
Explanation of what happens after a report is made	Description of the process following the mandated reporter's conversation with a child protective services intake worker (i.e., potential outcomes of the report, timeline for next steps)	40 (82%)	AL, AK, AR, CA, CO, CT, DC, FL, GA, ID, IL, IN, IA, KS, ME, MD, MA, MI, MN, MO, MT, NV, NH, NJ, NM, NY, NC, ND, OK, OR, PA, SD, TN, TX, VT, VA, WA, WV, WI, iLookOut

Additional Resources

A smattering of MR trainings provided additional resources to promote ongoing learning, such as *PDF handouts* that reinforced important learning points (n=10), *reporting worksheets* to help guide users through the process of making a report of suspected child abuse (n=4), and *links to online resources* such as government websites and state laws (n=21). Despite strong evidence that learning requires reinforcement—ideally using spaced retrieval and spaced practice (Burns & Gurung, 2020; Karpicke & Bauernschmidt, 2011)—only iLookOut provided additional learning exercises designed to reinforce and augment its MR training. These micro-learning exercises comprise iLookOut's Advanced Trainings

1 and 2 (Kapp et al., 2020; Levi et al., 2019), which learners can access following the completion of the iLookOut Core Training. For a comprehensive inventory of characteristics of the 49 mandated reporter trainings reviewed, see Table 3 and also <https://webgis.pop.psu.edu/iLookOut/>.

Gamification and Engagement

In the context of education and learning, gamification (e.g., storylines with hidden information, badges, points, avatars, matching exercises) has been shown to improve learner engagement and motivation, and contribute to higher learning outcomes (Dichey & Dicheva, 2017; Jang et al., 2015; Mohammed et al., 2018). Of the

Mandated Reporter Training

49 MR trainings in this comparative assessment, only nine made use of gamified activities. From an experiential standpoint, trainings were categorized based on their overall level of engagement. Of the

49 MR trainings reviewed, six were evaluated as Limited, 33 Basic, nine Moderate, and one Advanced. As detailed in Table 4, what distinguished more engaging MR trainings was their scope and use of audio-visual content and interactivity.

Table 4. Level of Engagement.

Level of engagement	Definition	States at this level	States
Advanced	Training includes multiple interactive* features, engaging multi-media formats, a wide array of resources, a scope that is considerably broader than just mandated reporting (e.g., trauma-informed care, mindfulness, critical thinking, support for families), both pre- and post-tests, and interactive feedback on knowledge test.	1 (2%)	iLookOut
Moderate	Training includes one or more interactive* features, requires participant engagement through frequent knowledge checks, may include a pre- or post-test, and includes information that goes beyond mandated reporter training.	9 (19%)	CA, GA, IN, KS, MD, MO, NM, UT, WI,
Basic	Training includes videos or audio-clips, a few minor interactive* features, and expanded information (typically as text) related to mandated reporting, such as legal requirements, signs of abuse, and prevention.	33 (67%)	AL, AK, AZ, AR, CO, CT, DC, FL, ID, IL, IA, LA, ME, MA, MI, MN, MT, NE, NH, NV, NY, NC, ND, OH, OR, PA, SC, SD, TX, VT, VA, WA, WV

Mandated Reporter Training

Level of engagement	Definition	States at this level	States
Limited	Training does not include interactive* features and presents information simply as text, a slideshow, and/or a single video.	6 (12%)	DE, HI, KY, NJ, OK, TN

Discussion

This comparative assessment identified significant variation in both content and delivery/functionality among 49 online (English language) mandated reporter (MR) trainings. Because all children deserve protection from abuse regardless of where they live, such variation raises concern over just how many mandated reporters in the U.S. have access to comprehensive preparation for recognizing and reporting suspected child abuse. This is particularly relevant if in-person MR training is not easily accessible, be it due to cost, timing, location, or other barriers.

iLookOut was designed to provide an evidence-based, standardized MR training that can be adapted to meet any given state's laws and policies. Included in this design is an emphasis on helping learners develop and apply critical thinking skills as they apply to suspected child abuse and, more generally, promoting child well-being. Whether it involves distinguishing poverty from neglect or raising awareness about cultural differences, we believe that well-designed MR trainings should include strategies for countering systemic racism and implicit bias. There was no ready metric for coding MR trainings on this goal, and so it was not evaluated in this comparative assessment.

Clearly, not all online MR trainings are created equal with regard to educating, engaging, and motivating mandated reporters. Perhaps the most glaring finding from this study is that 37 MR trainings failed to include both a pre- and post-test, and 17 MR trainings contain neither. This means that it is not possible to measure whether any one of these

37 state-approved MR trainings has any effect on mandated reporters' knowledge about child abuse and its reporting. In fact, a subsequent literature review found no published evaluation or outcomes studies for any of the 48 online MR trainings that we compared with iLookOut's Core Training. Further, for those MR trainings that had a pre- or post-test, we found no evidence that any of these other than iLookOut (Levi, et al., in press; Panlilio et al., in press) had validated their measures—as is needed to ensure that question items are truly evaluating their intended construct. So, too, no MR training other than iLookOut employed gamification or spaced retrieval/practice to promote learner engagement. To the extent that we want to both engage mandated reporters and optimize knowledge gain and retention, online trainings should take advantage of evidence-based practices shown to improve knowledge, change attitudes, and (ideally) affect people's actual behavior.

Limitations

Despite the breadth of our examination, there are several limitations to the present study. First, because only English versions of online MR training were reviewed, the content and functionality of MR trainings in other languages were not assessed. Second, because we did not continue searching MR training programs after identifying a state-sponsored training, it is possible that higher quality MR trainings exist in those states that had a state-sponsored MR training. Third, because there are no established criteria of what components should be included in MR training, the list of 40 components used to code the trainings may be neither

Mandated Reporter Training

comprehensive nor quintessential. Finally, because there is no existing standard for evaluating a MR training's level of engagement, there may be more appropriate criteria than were used in this study.

Conclusion

The findings of this comparative assessment show major state-to-state differences in the content and delivery/functionality of state-approved mandated reporter trainings. Because, as has been noted in prior research (Mathews & Kenny, 2008), there are non-trivial differences between the states in terms of policy, legal definitions, and reporting requirements, some amount of variability is to be expected. That being said, if it is worth investing the time, energy, and resources to educate mandated reporters, it is certainly worth ensuring that key concepts and strategies for protecting children are conveyed effectively. Otherwise, there may be little reason to believe that such training will actually help mandated reporters protect children.

Recommendations

Based on the study findings, we recommend the following suggestions to practitioners and policy makers:

- Establish national standards for what should be covered in MR training.
- Establish a national standard for rating the quality of online MR training, including criteria for what counts as an evidence-based training.
- Increase funding to develop MR training that incorporates best practices for online learning (e.g., spaced retrieval/practice) as well as innovations (e.g., gamification) that make such training genuinely engaging (and thereby more effective) for users.
- Encourage states to accept only MR training that is evidence based.
- Develop online evidence-based training that helps parents and other caregivers better understand the parameters of child abuse and its reporting.



About the Authors

Danielle Somerville, BS, is a research specialist who has over five years of experience in the area of child maltreatment. While pursuing her undergraduate degree at Penn State, she was involved in research studying the impact of child maltreatment on cognitive development. She has worked with the iLookOut project since 2019.

Sarah Dore, MPA, is Director of Learning and Curriculum for the iLookout project and has an extensive background in child welfare and program management. She has worked on the iLookOut project since 2016. She has been integral in developing its learning content and integrating gamification into learning activities.

Katherine L. Humphreys, PhD, EdM, is Assistant Professor in the Department of Psychology and Human Development, Vanderbilt University, and an infant mental health psychologist and researcher. Her lab examines the effects of early adversity (including abuse and neglect) on brain and behavior, the transition to parenthood, and developmental psychopathology. She has been affiliated with the iLookOut project since 2018.

Nicole Verdiglione, MA, is a research specialist who has extensive experience working on research projects within the Humanities and Pediatric Departments at the Penn State College of Medicine and has worked on the iLookOut project since 2014. Ms. Verdiglione has experience with multiple NIH funded research projects and has expertise assisting with study design, implementation, and data management as well as with early childhood education.

Shouhao Zhou, PhD, is Assistant Professor of the Department of Public Health Sciences at Pennsylvania State University. Dr. Zhou's research interests include clinical trial design, Bayesian hierarchical modeling, meta-analysis, and statistical model selection.

Robert Hamm, PhD, is Professor with the Department of Family and Preventive Medicine, University of Oklahoma Health Sciences Center. Dr. Hamm studies rational and erroneous decision-making, referring to the standards of decision theory and empirical accuracy and descriptions of how people use the information available to them. He has been affiliated with the iLookOut project since 2014.

Richard Fiene, PhD, a research psychologist, has spent his professional career in improving the quality of child care throughout the United States and internationally. Dr. Fiene has done extensive research and publishing on the key components in improving child care quality through an early childhood program quality indicator model of training, technical assistance, quality rating and improvement systems, professional development, mentoring, licensing, risk assessment, differential program monitoring, key indicators, and accreditation. He has been affiliated with the iLookOut project since 2014.

Karl Kapp, EdD, is Professor of Instructional Technology at Bloomsburg University in Bloomsburg, Pennsylvania, where he teaches instructional game design, gamification classes and online learning design. Dr. Kapp is the Director of Bloomsburg's Institute for Interactive Technologies and is recognized internationally as an expert in the application of games, game-thinking, and gamification to learning. He has been affiliated with the iLookOut project since 2014.

Benjamin Levi, MD, PhD, FAACP, is Professor of Pediatrics and Humanities, Penn State College of Medicine and Penn State Children's Hospital. Dr. Levi is a philosopher and practicing pediatrician whose research focuses on various aspects of decision-making. He also teaches bioethics, conducts clinical ethical consultations, and was the Founding Director of Penn State's Center for the Protection of Children. He is the Founder and Director of the iLookOut project. Contact: BHLEVI@psu.edu

References

Comparison of State Online Mandated Reporter Trainings

- Ayling, N. J., Walsh, K., & Williams, K. E. (2019). Factors influencing early childhood education and care educators' reporting of child abuse and neglect. *Australasian Journal of Early Childhood*, 45(1), 95–108. <https://doi.org/10.1177/1836939119885307>
- Burns, K. C., & Gurung, R. A. R. (2020). A longitudinal multisite study of the efficacy of retrieval and spaced practice in introductory psychology. *Scholarship of Teaching and Learning in Psychology*. <https://psycnet.apa.org/record/2020-59641-001>
- CARE House of Oakland County. (2020). *Online mandated reporter training*. Retrieved October 20, 2020, from <https://www.carehouse.org/mandated-reporter-training/>
- Chen, Y., Fetzter, S., Lin, C., Huang, J., & Feng, J. (2013). Healthcare professionals' priorities for child abuse educational programming: A Delphi study. *Children and Youth Services Review*, 35(1), 168–173. <https://doi.org/10.1016/j.childyouth.2012.09.024>
- Child Welfare Information Gateway. (2019). *Mandatory reporters of child abuse and neglect*. U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau. Retrieved April 4, 2021, at <https://www.childwelfare.gov/topics/systemwide/laws-policies/statutes/manda/>
- Damashek, A., Balachova, T., & Bonner, B. (2011). Training competent psychologists in the field of child maltreatment. *Journal of Clinical Psychology*, 67(7), 752–757. <https://doi.org/10.1002/jclp.20803>
- Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed, and what remains uncertain: A critical review. *International Journal of Educational Technology in Higher Education*, 14(1), 1–36. <https://doi.org/10.1186/s41239-017-0042-5>
- Errington, E. P. (2008). Exploring real-world scenarios as vehicles for authentic learning. *The International Journal of Interdisciplinary Social Sciences: Annual Review*, 3(5), 1–6. <https://doi.org/10.18848/1833-1882/cgp/v03i05/52619>
- Humphreys, K. L., Piersiak, H. A., Panlilio, C. C., Lehman, E. B., Verdiglione, N., Dore, S., & Levi, B. (n.d.). Evaluating changes in knowledge and attitudes following child abuse mandated reporter training: A randomized controlled trial of iLookOut. (under review).
- Jang, J., Park, J. J., & Yi, M. Y. (2015). Gamification of online learning. In C. Conati, N. Heffernan, A. Mitrovic, & M. Verdejo (Eds.), *Artificial intelligence in education (AIED)*. *Lecture Notes in Computer Science*, 9112. https://doi.org/10.1007/978-3-319-19773-9_82
- John, R. S., Salas-Wright, C. P., Amodeo, M., Chassler, D., & Alford, D. P. (2020). Faculty education in addiction training (FEAT): Evaluating an online training program for multidisciplinary health professions educators. *Substance Abuse*, 41(3), 292–296. <https://doi.org/10.1080/08897077.2020.1783739>

References, cont.

- Kapp, K., Dore, S., Fiene, R., Bennett, E., Grable, B., Hamm, R., & Levi, B. (2020, April). Cognitive mapping for iLookOut for Child Abuse: An online training program for early childhood professionals. *Online Journal of Distance Education and e-Learning*, 8(2), 80–89. <http://tojdel.net/?pid=showissue&issueid=236>
- Karpicke, J., & Bauernschmidt, A. (2011). Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1250–1257. http://learninglab.psych.purdue.edu/downloads/2011_Karpicke_Bauernschmidt_JEPLMC.pdf
- Kenny, M. C. (2015). Training in reporting of child maltreatment: Where we are and where we need to go. *Mandatory Reporting Laws and the Identification of Severe Child Abuse and Neglect Child Maltreatment*, 327–346. https://doi.org/10.1007/978-94-017-9685-9_16
- Kenny, M. C., & Abreu, R. L. (2015). Training mental health professionals in child sexual abuse: Curricular guidelines. *Journal of Child Sexual Abuse*, 24(5), 272–591. <https://doi.org/10.1080/10538712.2015.1042185>
- Kenny, M. C., Lopez-Griman, A. M., & Donohue, B. (2016). Development and initial evaluation of a cost-effective, Internet-based program to assist professionals in reporting suspected child maltreatment. *Journal of Child & Adolescent Trauma*, 10(4), 385–393. <https://doi.org/10.1007/s40653-016-0110-3>
- Levi, B. H., Belser, A., Kapp, K., Verdiglione, N., Mincemoyer, C., Dore, S., Keat, J. & Fiene, R. (2019). iLookOut for Child Abuse: Conceptual and practical considerations in creating an online learning programme to engage learners and promote behaviour change. *Early Child Development and Care*, 1–10. <https://doi.org/10.1080/03004430.2019.1626374>
- Levi, B. H., Mundy, M., Palm, C., Verdiglione, N., Fiene, R., & Mincemoyer, C. (In press). Developing an interactive online learning program on child abuse and its reporting. *Journal of Educators Online*.
- Levi, B., & Loeben, G. (2004). Index of suspicion: Feeling not believing. *Theoretical Medicine and Bioethics*, 25(4), 277–310. <https://doi.org/10.1007/s11017-004-3136-8>
- Mathews, B., & Kenny, M. C. (2008). Mandatory reporting legislation in the United States, Canada, and Australia: A cross-jurisdictional review of key features, differences, and issues. *Child Maltreatment*, 13(1), 50–63. <https://doi.org/10.1177/1077559507310613>
- Mathews, B., Yang, C., Lehman, E. B., Mincemoyer, C., Verdiglione, N., & Levi, B. H. (2017). Educating early childhood care and education providers to improve knowledge and attitudes about reporting child maltreatment: A randomized controlled trial. *PLoS ONE*, 12(5). <https://doi.org/10.1371/journal.pone.0177777>
- Mohammed, G. S., Wakil, K., & Nawroly, S. S. (2018). The effectiveness of microlearning to improve students' learning ability. *International Journal of Educational Research Review*, 3(3), 32–38. <https://doi.org/10.24331/ijere.415824>

Mandated Reporter Training

References, cont.

- Panlilio, C., Famularo, L., Masters, J., Dore, S., Verdiglione, N., Yang, C., Lehman, E., Hamm, R. M., Fiene, F., Kapp, K., & Levi, B. H. (In press). Integrating validity evidence to revise a child abuse knowledge test for early childhood education providers: A mixed methods approach. *American Journal of Evaluation*.
- SafeSchools. (2020, October 9). *Child abuse: Mandatory reporting*. Retrieved October 20, 2020, from <https://www.safeschools.com>
- Scott, M., Feldman, B. N., & Underwood, M. (2016). Delivering professional development in suicide prevention. *Pedagogy in Health Promotion*, 2(4), 266–275. <https://doi.org/10.1177/2373379916658667>
- Shendell, D. G., Apostolico, A. A., Milich, L. J., Patti, A. A., & Kelly, S. W. (2016). Comparing efficacy of online and in-person versions of a training on U.S. federal wage and hour, child labor laws, and hazardous occupations orders for secondary school professionals. *Frontiers in Public Health*, 4. <https://doi.org/10.3389/fpubh.2016.00075>
- Yang, C., Panlilio, C., Verdiglione, N., Lehman, E. B., Hamm, R. M., Fiene, R., Dore, S. Bard, D., Grable, B., & Levi, B. (2020). Generalizing findings from a randomized controlled trial to a real-world study of the iLookOut, an online education program to improve early childhood care and education providers' knowledge and attitudes about reporting child maltreatment. *PLoS ONE*, 15(1). <https://doi.org/10.1371/journal.pone.0227398>





Published in final edited form as:

Online J Distance Educ Elearn. 2020 April ; 8(2): 80–89.

COGNITIVE MAPPING FOR ILOOKOUT FOR CHILD ABUSE: AN ONLINE TRAINING PROGRAM FOR EARLY CHILDHOOD PROFESSIONALS

Karl M. Kapp, Ed.D.*,

Bloomsburg University, Department of Instructional Technology, 203 Sutliff Hall, Bloomsburg, PA 17815

Sarah Dore,

iLookOut for Child Abuse| Learning & Development Director, Department of Humanities, Penn State College of Medicine

Richard Fiene, Ph.D. [Affiliate Professor]

Edna Bennett Pierce Prevention Research Center, Professor of HDFS & Psychology (retired), The Pennsylvania State University

Breanna Grable,

Carlomagno Panlilio

Penn State University

Robert M. Hamm,

University of Oklahoma

Chengwu Yang

New York University

Erik B. Lehman,

Claudia Mincemoyer,

Nicole Verdiglione,

Benjamin H. Levi

Penn State University

Abstract

This article delineates the theory and framework for an innovative child abuse training program for mandated reporters called ‘iLookOut’. iLookOut is an online learning delivery system that utilizes mastery learning and self-determination theory in the Core Training program, along with spaced retrieval and retrieval practice in a follow-up micro-learning program that reinforces learning from the Core Training. A cognitive mapping model provides the structure for documenting and organizing the learning content in both the Core training and the follow-up micro-learning program. The article provides a conceptual framework for designing and implementing effective and efficient online learning programs.

*Primary Contact kkapp@bloomu.edu & karlkapp@gmail.com.

Keywords

distance learning; cognitive mapping; learning; engagement; online learning

INTRODUCTION

There are relatively few studies of Early Care & Education (ECE) professionals' child abuse reporting practices, particularly compared with published studies on other mandated reporters who have high levels of contact with children (e.g., teachers, nurses, doctors). In those that have been done, ECE professionals are less likely to have ever reported child maltreatment compared to other professionals who work with children (Zellman & Bell, 1990); this was due in part to the minimal education ECE professionals received about child abuse and what regarding the level of concern and/or circumstances warrant reporting (Alvarez, Kenny, Donohue, & Carpin, 2004; Kenny, 2007; Carter, Bannon, Limbert, Docherty, & Barlow, 2006).

The small body of quantitative and qualitative research examining reporting experiences of ECEs (including kindergarten and pre-school teachers (Sundell, 1997; Bishop, Lunn, & Johnson, 2002) has revealed high levels of uncertainty about the decision to report, perceived "conflicts of loyalty," (Svensson & Janson, 2008) and complexities that, taken together, have caused some ECE professionals to feel as if they are "dancing on the edge." (Feng, Chen, Wilk, Yang & Fetzer, 2009).

ECE professionals report wanting to preserve relationships with families and avoid causing harm, but at the same time meet their legal, professional, and ethical responsibilities. One consequence of such uncertainty and conflict is report latency, with one study finding an average time of 14 months between ECE professionals having suspicion and making a report (Sundell, 1997) -- a situation that, if left unaddressed, risks dire outcomes for many young children.

The emergence of the internet and other technological advances provide a mechanism for educating ECE professionals about child abuse and its reporting, even in remote areas. But evidence-based curricula are lacking. This paper describes a methodology for crafting an integrated curriculum that meets the needs of ECE professionals and can be delivered via online learning modalities.

The Need

Every day, thousands of children experience one form of child abuse or another at the hands of a parent or other caregiver. Annually, there are 680,000 confirmed cases in the United States, and research provides strong evidence that this number is likely much higher (Finkelhor, Turner, Ormrod, & Hamby, 2010; Hussey, Chang, & Kotch, 2006, Kohl, Jonson-Reid, & Drake, 2009; Stoltenborgh, Bakermans-Kranenburg, van Ijzendoorn, & Alink, 2013; Stoltenborgh, van Ijzendoorn, Euser, & Bakermans-Kranenburg, 2011; Stoltenborgh, Bakermans-Kranenburg, Alink, & van Ijzendoorn, 2012; Stoltenborgh,

Bakermans-Kranenburg, & van Ijzendoorn, 2013; Sedlak, Mettenburg, Basena, Peta, McPherson, & Greene, 2010).

The youngest children –those under 5 years of age– are more likely to be victims of all forms of child abuse other than sexual abuse. Because they are more vulnerable, these children experience serious injuries –including death– at much higher rates than older children. (U.S. Department of Health & Human Services, 2018). However, not all consequences of abuse are as immediate as broken bones and black eyes. Each of the various forms of abuse can have profound effects on children’s physical, psychological, developmental, and overall emotional well-being –with conclusive evidence of the strong relationship between child maltreatment and subsequent anxiety, depression, substance use, intimate partner and family violence, as well as heart disease, strokes, and cancer (Norman, Byambaa, De, Butchart, Scott, & Vos, 2012, Flaherty et al., 2013; Jonson-Reid, Kohl, & Drake, 2012; Mills et al., 2011; Danese & McEwen, 2012; Shonkoff & Garner, 2012; Hadland et al., 2015).

In the face of these staggering facts, it is perhaps surprising that ECE professionals –who provide care to 8–12 million children in the U.S.– (Laughlin, 2013; Laughlin, 2006) report only about half of 1% of confirmed cases of child abuse. (U.S. Department of Health & Human Services, 2018).

Because ECE professionals interact with so many young children on a daily basis, they are in a unique position to identify and respond to suspected child abuse. They may be the only people outside of a child’s immediate family to have extended opportunities to discern red flags and/or subtle signs of abuse that might lead to early detection, the potential to help prevent patterns of abuse from taking hold, and the opportunity to support over-worked and overwhelmed families before abuse occurs. With the right preparation and resources, ECE professionals can also serve as key supports for children and families who are struggling (Dinehart, Katz, Manfra, & Ullery, 2013). Yet protecting children is far from straight-forward or low stress task –to the extent that some ECE professionals have identified “reporting possible abuse” as the most troubling ethical issue they face in their workplace (Clyde & Rodd, 1989; Feeney & Sysko, 1986).

Two-Phased Approach

The *iLookOut for Child Abuse* learning program (*iLookOut*) has two distinct phases. The first is the Core Training, which uses a video-based storyline, experiential learning theory, and practice modules to provide ECE professionals a strong, standardized foundation for understanding what child abuse is, what to look for, and what to do (and not do).

The second phase provides Advanced Training using spaced practice and spaced retrieval to augment the original instruction and provide continuing reinforcement delivered via email or a smart phone application. We have used the term “pinging” for as shorthand to represent the combination of spaced practice and spaced retrieval. We chose the term “pinging” because we think of the process as similar to the sound waves that submarines send out to gauge progress and location. In our case, rather than sound waves, our Advanced Training pings offer micro-learning opportunities, reminders about what to look for in cases of child abuse

and feedback on progress toward completing the training. Such continual reinforcement related to child abuse encourages ECE professionals to re-process, synthesize and (most importantly) apply what they have learned about child abuse and what can be done to help protect children and promote their well-being. Because *iLookOut*'s pinging is iterative, and dispatches brief messages over time, this micro-learning can be tailored to the availability and needs of individual learners.

This paper describes *iLookOut*'s two phases, and explains how the distance learning curriculum that is now deployed to ECE professionals across Maine and Pennsylvania (Core Training only/Phase I) can serve as a model for others looking to deliver trainings and associated resources and reinforce learning over broad geographic regions. In this context, "Core Training" refers to the initial 3-hour *iLookOut* learning program, while "Advanced Training" refers to the micro-learning activities sent out as pings to smart-phones or other mobile technologies.

Core Training

Given the broad and dispersed population of ECE professionals in Maine, and the acute need for mandated reporter training, we sought to create an online program that could be delivered to ECE professionals where they worked. The *iLookOut* Core Training uses a video-based storyline and game-based techniques to more effectively engage ECE professionals, along with pre/post-testing to measure knowledge, attitudes, and satisfaction (Levi et al., 2019)

Like other online programs, *iLookOut*'s Core Training provides ready, low-cost access to multi-media learning 24/7, and can be paused/resumed as desired. *iLookOut*'s content is written at an 8th grade level, and provides standardized education for a workforce known for wide variability in entry level training, skill-sets, work environment, and professional development opportunities.

Mastery Learning—The *iLookOut*'s Core Training was designed to ensure that all ECE professionals could master basic information about how to identify signs of child abuse. This mastery learning philosophy recognizes that under appropriate instructional conditions virtually all learners can master what is being taught (Block & Burns, 1976; Bloom, 1971). *iLookOut*'s Core Training creates those conditions by organizing the curriculum into discrete units; providing interactive instructional activities along with various didactic exercises and resource handouts; embedding assessments within these activities; and requiring learners to master the learning objectives before moving to the next discrete unit. (Bloom, 1971; Melton, 2008).

Mastery learning is an approach that recognizes that aptitude for learning may be more closely linked to an individual's perseverance and time spent than to any notion of "ability" (Bloom, 1971; Melton, 2008). As with criterion-referenced tests, which assess the performance of each test-taker without regard to the performance of others (Shrock & Coscarelli, 2007), there is no limit to the number of ECE professionals who can excel in completing the *iLookOut* curriculum.

Self-Determination Theory—Motivationally, *iLookOut*'s Core Training program is based on Self-Determination Theory (SDT). This macro-theory which has been used to explain human motivation in many endeavors (including sports, healthcare, religion, work, and education) posits that human beings primarily perform tasks/activities because of an internal drive rather than some externally driven theory of operant conditioning (Ryan & Deci, 2000a; Ryan & Deci, 2000b). Among other things, SDT has helped to identify factors that either facilitate or undermine human motivation. For example, one of SDT's sub-theories, cognitive valuation theory, proposes that events and conditions that enhance a person's sense of autonomy and competence intrinsically support motivation, while factors that diminish perceived autonomy or competence undermine intrinsic motivation.

In the context of SDT, autonomy involves a person feeling that they are in control of their actions, and can influence the outcome of those actions. To help foster this sense of autonomy, the *iLookOut* Core Training provides ECE professionals with opportunities to make meaningful choices in response to the scenarios playing out in the video-based storyline –particularly with regard to possible signs of child abuse.

Another key aspect of SDT involves the human drive to take on challenges and achieve a sense of mastery. SDT describes this as striving for competence, and posits that factors that enhance an individual's ability to experience competence (e.g., opportunities to acquire new skills or overcome challenges) are intrinsically motivating. *iLookOut*'s Core Training provides ECE professionals many such opportunities to be challenged, to demonstrate mastery, and to earn digital badges that offer visible acknowledgement of their achievements.

iLookOut also incorporates SDT's third major element, relatedness –the experience of feeling meaningfully connected with others. *iLookOut*'s Core Training does this by helping ECE professionals identify with being part of a profession and community that is united in its goal of promoting children's well-being.

Pinging and Advanced Training

Despite the many strengths of *iLookOut*'s Core Training, it is well established that gains in knowledge are quickly lost unless they are somehow reinforced (Murre & Dros, 2015). Because decades of research show that spaced practice and spaced retrieval optimize learning (Ausubel, & Youssef, 1965; Caple, 1996; Kerfoot, 2010), the *iLookOut* programs were designed to include reminders of interactive micro-learning activities that are sent to learner's smart-phones and emails. These pings serve to reinforce and augment the concepts that were originally taught in the Core training.

Spaced Retrieval—Broadly speaking, the concept of *spaced retrieval* involves providing learners with course content spaced over time. Spaced retrieval has been shown to be an effective tool for aiding student retention (Carpenter & DeLosh, 2005), and has been more widely promoted through computer-based adaptive instructional models such as ALEKS (Doignon & Falmagne, 1985) and LearnSmart (McGraw-Hill, 2013). In contrast to the standard practice of asking learners to digest large amounts of content all at one time, spaced retrieval avoids learner fatigue, as well as setting unrealistic expectations. For *iLookOut*'s

Advanced Training phase, spaced retrieval also allows learners time to process and reflect on new information at multiple points in time, rather than simply moving on.

Spaced retrieval helps learners retain access to memorized information over long periods of time because the spacing promotes deeper processing of the learned material. Ideally, the time between the learning events is greater than 24 hours, but shorter times have also been found to be effective. As long as eight years after an initial training, learners who engaged in spaced retrieval exercises showed better retention than those whose learning was more concentrated time period (Clark & Mayer, 2011).

iLookOut's Advanced Training sends participants weekly pings, each constituting a learning module that includes an activity (reading, game, video, etc.) along with various question-items. Learners must complete the modules in succession, and may review their content after completion; but learners may not proceed to the next module until its predetermined release date. Successful completion of a set number of modules ultimately earns learners a badge.

Retrieval Practice—By design, our use of retrieval practice requires ECE professionals to recall or retrieve information they have learned, and complete both “knowledge checks” and In-Practice exercises that provide opportunities to apply newly acquired knowledge. The benefits of retrieval practice are well-established across diverse groups (Larsen et al., 2009). But the advent of computer technology has added the ability to not only record learner responses and performance, but also standardize and formally integrate spaced practice into learning curricula.

Retrieval practice improves recall performance in part because the act of retrieving information from memory actually strengthens the existing memory trace, and often creates additional retrieval routes (Dobson, 2013). Because these changes increase the probability of successful retrieval in the future (Roediger & Butler, 2013), retrieval practice can significantly enhance long-term retention of what ECE professionals have learned.

Despite the known efficacy of spaced practice and retrieval practice, it was not obvious how best to apply them vis-à-vis *iLookOut*'s overall curriculum. To develop a systematic approach for doing so –i.e., to determine the appropriate sequencing and content for Phase 2 of *iLookOut*—we adopted a “Cognitive Mapping” approach to design a comprehensive pinging curriculum.

Cognitive Mapping

Cognitive Mapping was first introduced in 1948 by educational psychologist, Edward Tolman (Tolman, 1948) to explain how rats learned the locations of rewards in a maze, and as such generated a practical model for mapping their environment. Cognitive mapping is now in wide use in many different venues (including health research (Stadler, et al, 2013) and engineering (Dixon & Lammi, 2014)), not only to identify and illustrate how key elements are (or should be) inter-related, but also to create strategies for integrating, measuring, and analyzing various factors and outcome.

Cognitive mapping helps explore learner characteristics that improve learning (e.g., self-regulated learning components of goal-directedness, motivation, goal feedback, etc.). One such characteristic involves outcome expectations (Schunk & Zimmerman, 2006). Rooted in Tolman's concept of field expectancies, a learner's outcome expectations are based in their ability to anticipate particular relationships between a (e.g., lightning) stimulus and a response (e.g., thunder). Such expectancies help people form cognitive maps, which are internal representations of these expectancies, along with a catalogue of actions that are more (or less) likely to help individuals attain their goals.

Cognitive maps are particularly important for latent learning –i.e., learning that occurs after the initial period of teaching/exploration/etc. Latent learning may occur at any number of points in time, but may be most pronounced when the learner realizes how what has been taught applies in real-world settings –which in the case of ECE professionals is likely to be when they are working with infants and toddlers.

According to a Social Cognitive framework, learners will act in a manner they believe is likely to be successful, and will adopt observational and behavioral frameworks that conduce to success. As such, *iLookOut's* cognitive map was designed to both 1) help our research team understand the relations between the various components of the Core Training and the subsequent pings, and 2) create a prototype of the internal model we hoped our learners would develop over the course of *iLookOut's* two phases. In this way, the cognitive map developed for *iLookOut* provides a template to help ECE professionals more effectively connect and integrate information, ways of observing/interpreting, and particular practices so as to optimize and take full advantage of latent learning.

By definition “cognitive maps” are mental or conceptual models, “thinking maps” that, like other forms of cartography, map territory. But here it is cognitive “territory” rather than geographic terrain that is being characterized. Sometimes, cognitive maps provide a linear progression of a concept, or the relationships between various factors. They also can serve a developmental purpose, by helping people (be they researchers, policy makers, teachers, or learners) develop a deeper understanding of how different elements are (or should be) related to one another.

There are many applications of cognitive maps. *Perceptual* applications tend to focus on a) inquiring and/or gathering information; b) noticing/attending to; or c) differentiating/distinguishing. *Cognitive* applications tend to focus on a. organizing data and/or finding patterns/relationships; b) interpreting/understanding data; c) analyzing data; d) troubleshooting/diagnosing; e) drawing conclusions; f) framing; g) illuminating; or h) estimating probability/confidence levels. *Evaluative* applications focus on a) assessing/judging; b) measuring effectiveness; c) predicting future success; d) assigning importance/priority; or e) providing feedback. *Volitional* applications include a) identifying desires; b) defining/clarifying purposes; or c) planning. *Behavioral* applications include a) performing/behaving; b) implementing a plan; c) communicating; d) learning; e) improving skills; or f) developing.

Of these, *iLookOut*'s cognitive map was developed for the following purposes: coordinating learning content (Core Training and Advanced Training), implementing a plan, predicting likelihood and degree of future success, providing feedback, monitoring and measuring progress, evaluating and assessing achievement, organizing data, and finding patterns. This allowed us to both fully align existing content, strategies, and outcomes, and also identify critical areas that were not sufficiently fleshed out or appropriately integrated. For *iLookOut*'s Advanced Training in particular, this involved 1) distinguishing the learning points for the different concepts; 2) demonstrating associations between various components; 3) mapping individuals' actual progress through different stages of learning; 4) clarifying the purposes of various components of the learning program; and 5) creating a framework for how learners will progress through various activities so as to develop their skills. Figure 1 provides a very simplified cognitive map showing how key elements from *iLookOut*'s Core Training and Advanced Training fit together to create a unified whole. Such integration is crucial because any misalignment could potentially confuse or demotivate learners, fail to leverage spaced retrieval/practice, and/or undermine latent learning.

Figure 2 provides more detailed mapping of the Core Training, showing where and how learning objectives manifest in the storyline, discrete learning modules, and post-training knowledge tests.

Figure 3 provides a more detailed mapping of a portion of the Advanced Training, corresponding to the learning content in Figure 2. In addition to listing the topic and content for each week's pings, this cognitive map shows the kind of activity and teaching modality learner will experience.

Conclusion

The mapping process is critical for ensuring that key learning objectives are aligned with both content and teaching modalities, and that concepts and educational activities are appropriately sequenced to support the developmental goals of *iLookOut*. This cognitive mapping also helps to ensure consistency, cohesiveness, and alignment of the Core and Advanced Trainings.

The use of cognitive mapping is not common in non-academic settings. Yet, as described in this paper, this process can have great value for developing conceptually rich and well-integrated training programs, particularly for those wishing to leverage the power of spaced retrieval and spaced practice. This can be particularly valuable for topics like child abuse that are both contextually nuanced and emotionally complex. As such, this description of *iLookOut* is presented as a prototype that other researchers, designers, and developers of curricula may wish to consider and improve upon.

References

- Al-Azawi R, Al-Faliti F, & Al-Blushi M (2016). Educational Gamification Vs. Game Based Learning: Comparative Study. *International Journal of Innovation, Management and Technology*, 131–136. doi:10.18178/ijimt.2016.7.4.659

- Alvarez KM, Kenny MC, Donohue B, & Carpin KM (2004). Why are professionals failing to initiate mandated reports of child maltreatment, and are there any empirically based training programs to assist professionals in the reporting process? *Aggression and Violent Behavior*, 9(5), 563–578. doi:10.1016/j.avb.2003.07.001
- Ausubel DP, & Youssef M (1965). The Effect of Spaced Repetition on Meaningful Retention. *The Journal of General Psychology*, 73(1), 147–150. doi:10.1080/00221309.1965.9711263 [PubMed: 14316956]
- Bishop A, Lunn P, & Johnson K (2002). I would just like to run away and hide, but I won't! Exploring Attitudes and Perceptions on Child Protection Issues with Early Years Teacher Trainees on the Threshold of Their Careers. *Westminster Studies in Education*, 25(2), 187–199. doi:10.1080/0140672022000029203
- Blair L, & Kapp KM (2012). The gamification of learning and instruction: Game-based methods and strategies for training and education. San Francisco, CA: Pfeiffer.
- Block JH, & Burns RB (1976). Mastery Learning. *Review of Research in Education*, 4, 3. doi:10.2307/1167112
- Block JH, & Bloom BS (1971). *Mastery learning: Theory and practice*. New York: Holt Rinehart.
- Caple C (1996). *The Effects of Spaced Practice and Spaced Review on Recall and Retention Using Computer Assisted Instruction*. MI: Ann Arbor.
- Calder J (June 2000). Beauty Lies in the Eye of the Beholder. *International Review of Research in Open and Distance Learning*, 1 (1), [Online]. doi:10.19173/irrodl.v1i1.6
- Carpenter SK, & Delosh EL (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19(5), 619–636. doi:10.1002/acp.1101.
- Carter YH, Bannon MJ, Limbert C, Docherty A, & Barlow J (2006). Improving child protection: A systematic review of training and procedural interventions. *Archives of Disease in Childhood*, 91(9), 740–743. doi:10.1136/adc.2005.092007. [PubMed: 16556613]
- Clark RC, & Mayer RE (2011). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (3rd ed.). New York, NY: Jossey-Bass/Pfeiffer.
- Clyde M, & Rodd J (1989). Professional Ethics: There's more to it than meets the eye! *Early Child Development and Care*, 53(1), 1–12. doi:10.1080/0300443890530101.
- Chugh R, Ledger S, & Shields R (2017) Curriculum Design for Distance Education in the Tertiary Sector. *Turkish Online Journal of Distance Education-TOJDE* April 2017 ISSN 1302-6488 Volume: 18 Number: 2 Article 1. *need article name
- Danese A, & McEwen BS (2012). Adverse childhood experiences, allostasis, allostatic load, and age-related disease. *Physiology & Behavior*, 106(1), 29–39. doi:10.1016/j.physbeh.2011.08.019. [PubMed: 21888923]
- Deterding S, Khaled R, Nacke LE, & Dixon D (2011, May). Gamification: Toward a definition. In *Proceedings of CHI 2011 Gamification Workshop*, (pp. 1–4). Vancouver, BC, Canada. Retrieved from <http://gamification-research.org/wp-content/uploads/2011/04/02-Deterding-Khaled-Nacke-Dixon.pdf>.
- Dinehart LH, Katz LF, Manfra L, & Ullery MA (2012). Providing Quality Early Care and Education to Young Children Who Experience Maltreatment: A Review of the Literature. *Early Childhood Education Journal*, 41(4), 283–290. doi:10.1007/s10643-012-0553-6.
- Dixon RA, & Lammi M (2014). Cognitive Mapping Techniques: Implications for Research in Engineering and Technology Education. *Journal of Technology Education*, 25(2). doi:10.21061/jte.v25i2.a.1
- Dobson JL (2013). Retrieval practice is an efficient method of enhancing the retention of anatomy and physiology information. *Advances in Physiology Education*, 37(2), 184–191. doi:10.1152/advan.00174.2012 [PubMed: 23728136]
- Doignon J, & Falmagne J (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, 23(2), 175–196. doi:10.1016/s0020-7373(85)80031-6
- Domínguez A, Saenz-De-Navarrete J, De-Marcos L, Fernández-Sanz L, Pagés C, & Martínez-Herráiz J (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, 63, 380–392. doi:10.1016/j.compedu.2012.12.020

- Murre JM, & Dros J (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *Plos One*,10(7). doi:10.1371/journal.pone.0120644
- Feeney S, & Sysko L (November, 1986). Professional ethics in early childhood education: Survey results. *Young Children*, 42(1), 15–20.
- Feng J, Huang T, & Wang C (2010). Kindergarten teachers' experience with reporting child abuse in Taiwan. *Child Abuse & Neglect*,34(2), 124–128. doi:10.1016/j.chiabu.2009.05.007 [PubMed: 20153054]
- Finkelhor D, Turner H, Ormrod R, & Hamby SL (2010). Trends in Childhood Violence and Abuse Exposure. *Archives of Pediatrics & Adolescent Medicine*,164(3), 238. doi:10.1001/archpediatrics.2009.283 [PubMed: 20194256]
- Flaherty EG, Thompson R, Dubowitz H, Harvey EM, English DJ, Proctor LJ, & Runyan DK (2013). Adverse Childhood Experiences and Child Health in Early Adolescence. *JAMA Pediatrics*,167(7), 622. doi:10.1001/jamapediatrics.2013.22 [PubMed: 23645114]
- Hadland S, Wood E, Dong H, Marshall B, Kerr T, Montaner J, & DeBeck K (2015). Suicide Attempts and Childhood Maltreatment Among Street Youth: A Prospective Cohort Study. *Pediatrics*,136(3). doi:10.1542/peds.2015-1108d
- Huffman S (2018, March 22). The Digital Divide Revisited: What Is Next? Retrieved from <https://www.questia.com/library/journal/1G1-531709941/the-digital-divide-revisited-what-is-next>
- Hussey JM, Chang JJ, & Kotch JB (2006). Child Maltreatment in the United States: Prevalence, Risk Factors, and Adolescent Health Consequences. *Pediatrics*,118(3), 933–942. doi:10.1542/peds.2005-2452 [PubMed: 16950983]
- Child and Family Services. (2018). Retrieved from <https://www.maine.gov/dhhs/ocfs/ec/occhs/infoparents.htm>
- Child and Adult Outcomes of Chronic Child Maltreatment. (2012). *Pediatrics*, 129(5). doi:10.1542/peds.2011-2529d
- Kapp KM (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. San Francisco: Pfeiffer.
- Kenny MC (2007). Web-based training in child maltreatment for future mandated reporters. *Child Abuse & Neglect*, 31(6), 671–678. doi:10.1016/j.chiabu.2006.12.008 [PubMed: 17590435]
- Kerfoot BP, Fu Y, Baker H, Connelly D, Ritchey ML, & Genega EM (2010). Online Spaced Education Generates Transfer and Improves Long-Term Retention of Diagnostic Skills: A Randomized Controlled Trial. *Journal of the American College of Surgeons*,211(3). doi:10.1016/j.jamcollsurg.2010.04.023
- Kohl PL, Jonson-Reid M, & Drake B (2009). Time to Leave Substantiation Behind. *Child Maltreatment*,14(1), 17–26. doi:10.1177/1077559508326030 [PubMed: 18971346]
- Larsen DP, Butler AC, & Iii HL (2009). Repeated testing improves long-term retention relative to repeated study: A randomised controlled trial. *Medical Education*,43(12), 1174–1181. doi:10.1111/j.1365-2923.2009.03518.x [PubMed: 19930508]
- Laughlin L (2010). Whos minding the kids?: Child care arrangements, Spring 2005/Summer 2006. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census.
- Laughlin L (2013). Who's Minding the Kids? Child Care Arrangements: Spring 2011. Retrieved from <https://www.census.gov/prod/2013pubs/p70-135.pdf>.
- Levi BH, Belser A, Kapp K, Verdiglione N, Mincemoyer C, Dore S, ... Fiene R (2019). Ilookout for child abuse: Conceptual and practical considerations in creating an online learning programme to engage learners and promote behaviour change. *Early Child Development and Care*,1–10. doi:10.1080/03004430.2019.1626374
- Mattingly MA *Child Youth Care Forum* (1995) 24: 379. 10.1007/BF02128529
- Melton KI (2008). Using Modified Mastery Assignments to Increase Learning in Business Statistics. *Decision Sciences Journal of Innovative Education*,6(2), 239–245. doi:10.1111/j.1540-4609.2008.00169.x
- Mills R, Alati R, O'Callaghan M, Najman JM, Williams GM, Bor W, & Statehearn L (2011). Child Abuse and Neglect and Cognitive Function at 14 Years of Age: Findings From a Birth Cohort. *Pediatrics*, 127(2), 4–10. doi:10.1542/peds.2009-3479d [PubMed: 21135010]

- Nacke LE, & Deterding S (2017). The maturing of gamification research. *Computers in Human Behavior*, 71, 450–454. doi:10.1016/j.chb.2016.11.062
- Norman RE, Byambaa M, De R, Butchart A, Scott J, & Vos T (2012). The Long-Term Health Consequences of Child Physical Abuse, Emotional Abuse, and Neglect: A Systematic Review and Meta-Analysis. *PLoS Medicine*, 9(11). doi:10.1371/journal.pmed.1001349
- Sailer M, Hense JU, Mayr SK, & Mandl H (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, 69, 371–380. doi:10.1016/j.chb.2016.12.033
- Iii HL, & Butler AC (n.d.). Retrieval Practice (Testing) Effect. *Encyclopedia of the Mind*. doi:10.4135/9781452257044.n246
- Ryan RM, & Deci EL (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67. doi:10.1006/ceps.1999.1020 [PubMed: 10620381]
- Ryan RM, & Deci EL (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. doi:10.1037/0003-066x.55.1.68 [PubMed: 11392867]
- Ryan RM, Rigby CS, & Przybylski A (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, 30(4), 344–360. doi:10.1007/s11031-006-9051-8
- Schunk DH, & Zimmerman BJ (n.d.). Competence and Control Beliefs: Distinguishing the Means and Ends. *Handbook of Educational Psychology*. doi:10.4324/9780203874790.ch16
- Sedlak AJ, Mettenburg J, Basena M, Petta I, Mcpherson K, Greene A, & Li S (2010). Fourth National Incidence Study of Child Abuse and Neglect (NIS-4): Report to Congress: Executive Summary. *PsycEXTRA Dataset*. doi:10.1037/e565022012-001
- Garner AS, Shonkoff JP, Siegel BS, Dobbins MI, Earls MF, Garner AS, ... Wood DL (2011). Early Childhood Adversity, Toxic Stress, and the Role of the Pediatrician: Translating Developmental Science Into Lifelong Health. *Pediatrics*, 129(1). doi:10.1542/peds.2011-2662
- Shrock SA, Coscarelli WC, & Sandbergh D (1998). Criterion-referenced test development: Technical and legal guidelines for corporate training. *Performance Improvement*, 37(8), 48–49. doi:10.1002/pfi.4140370812
- Stadler J, Dugmore C, Venables E, Macphail C, & Delany-Moretlwe S (2013). Cognitive mapping: Using local knowledge for planning health research. *BMC Medical Research Methodology*, 13(1). doi:10.1186/1471-2288-13-96
- Stoltenborgh M, Bakermans-Kranenburg MJ, Alink LR, & Ijzendoorn MH (2012). The Universality of Childhood Emotional Abuse: A Meta-Analysis of Worldwide Prevalence. *Journal of Aggression, Maltreatment & Trauma*, 21(8), 870–890. doi:10.1080/10926771.2012.708014
- Stoltenborgh M, Bakermans-Kranenburg MJ, & Ijzendoorn MH (2012). The neglect of child neglect: A meta-analytic review of the prevalence of neglect. *Social Psychiatry and Psychiatric Epidemiology*, 48(3), 345–355. doi:10.1007/s00127-012-0549-y [PubMed: 22797133]
- Stoltenborgh M, Bakermans-Kranenburg MJ, Ijzendoorn MH, & Alink LR (2013). Cultural–geographical differences in the occurrence of child physical abuse? A meta-analysis of global prevalence. *International Journal of Psychology*, 48(2), 81–94. doi:10.1080/00207594.2012.697165 [PubMed: 23597008]
- Stoltenborgh M, Ijzendoorn MH, Euser EM, & Bakermans-Kranenburg MJ (2011). A Global Perspective on Child Sexual Abuse: Meta-Analysis of Prevalence Around the World. *Child Maltreatment*, 16(2), 79–101. doi:10.1177/1077559511403920 [PubMed: 21511741]
- Sundell K (1997). Child-care personnels failure to report child maltreatment: Some Swedish evidence. *Child Abuse & Neglect*, 21(1), 93–105. doi:10.1016/s0145-2134(96)00133-0 [PubMed: 9023025]
- Svensson B, & Janson S (2008). Suspected Child Maltreatment: Preschool Staff in a Conflict of Loyalty. *Early Childhood Education Journal*, 36(1), 25–31. doi:10.1007/s10643-008-0248-1
- Tolman EC (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189–208. doi:10.1037/h0061626 [PubMed: 18870876]

- U.S. Department of Health and Human Services, Administration for Children and Families
Administration on Children, Youth and Families: Children's Bureau. (2018). Child Maltreatment
2016. Retrieved June 7, 2019 from <https://www.acf.hhs.gov/cb/resource/child-maltreatment-2016>.
- Werbach K, & Hunter D (2012). For the win: How game thinking can revolutionize your business.
Philadelphia: Wharton Digital Press.
- Winters D, & Latham GP (1996). The Effect of Learning versus Outcome Goals on
a Simple versus a Complex Task. *Group & Organization Management*, 21(2), 236–250.
doi:10.1177/1059601196212007
- Zellman G, & Bell RM (1990). The Role of Professional Background, Case Characteristics, and
Protective Agency Response in Mandated Child Abuse Reporting. Rand Corporation.

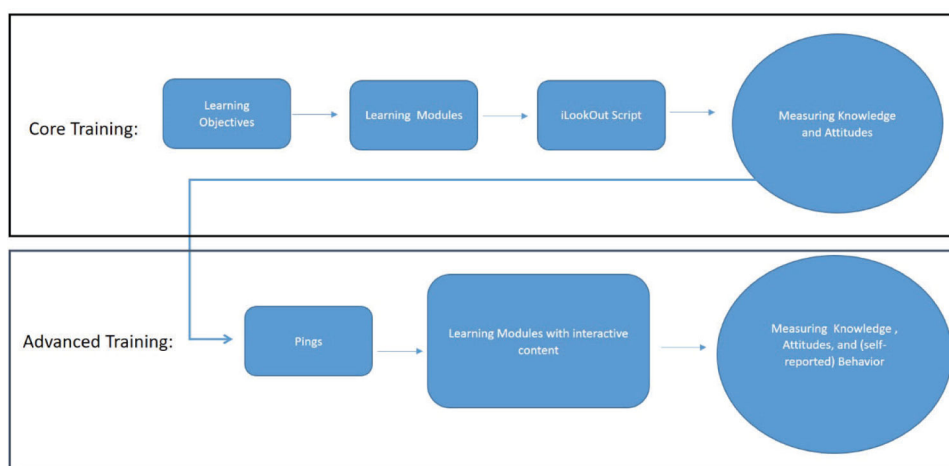


Figure 1.
Simplified Overview Cognitive Map

Learning Objectives (LO) of LookOut	Script Intervention & Learning Modules (LM)	Learning Module Type	Knowledge Test (KT) Conceptual Areas	Attitudinal Items
LO1: Understand and recognize possible child abuse, including: LO1.1: What does and does not constitute child abuse LO1.2: How and where abuse occurs LO1.3: Risk factors for abuse	(LM-1) Types of Abuse	Slide	KT1: Actions by adults that might constitute child abuse KT1.2: Places duct tape over a child's mouth as a form of punishment (Y/N/U) [LM-18] KT1.3: Causes any kind of physical injury to a child (Y/N/U) [LM-6] KT1.4: Causes a child substantial pain from disciplining him/her (Y/N/U) [LM-18] KT1.5: Impairs a child's physical functioning from discipline him/her (Y/N/U) [LM-6] KT1.6: Forcefully slaps a child under one year of age (Y/N/U) [LM-18] KT1.7: Physically abuses his/her partner (aka domestic violence) in the presence of a child (Y/N/U) [LM-16] KT1.8: Often has body odor, wears clothes that don't fit, and is inappropriately dressed for the weather (Y/N/U) [LM-5] KT1.9: Is always hungry, and often found stuffing food in his/her pockets to take home (Y/N/U) [LM-7]	Do you feel confident in your ability to identify signs of child abuse/neglect? (Self-Efficacy #1)
	(LM-6) Types of Child Abuse	Slide and Resource File Handout #3: Types of Child Abuse		
	(LM-7) "Possible Signs of Abuse"	Multiple Choice		
	(LM-16) "Intimate Partner Violence"	True/False		
	(LM-18) "Reporting Requirements"	Multiple Choice (Y/N/U)		
	(LM-9) Facts About Abuse	Resource File Handout #4: Facts About Abuse		
	(LM-8) "Thinking Carefully About Abuse"	Multiple Choice		

Figure 2.
Sample Cognitive Map for Core Training

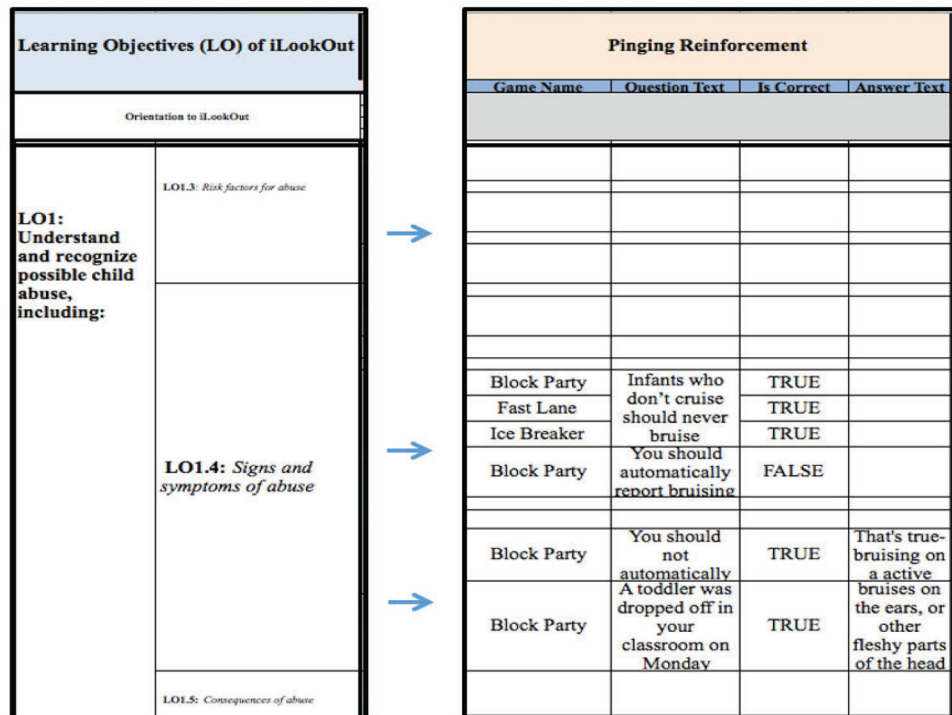


Figure 3.
Sample Cognitive Map for Advanced Training

AN INTERACTIVE ONLINE LEARNING PROGRAM ON CHILD ABUSE AND ITS REPORTING

Benjamin H Levi, Penn State College of Medicine
Mandy Mundy, Network of Victim Assistance
Cathleen Palm, Center for Children's Justice
Nicole Verdiglione, Penn State College of Medicine
Richard Fiene, Research Institute for Key Indicators
Claudia Mincemoyer, Penn State University

ABSTRACT

This article describes the content and pedagogical foundations of iLookOut for Child Abuse, an interactive, online learning program that was designed for early childhood professionals and others who provide childcare to young children. It also describes how an online intervention can address a complex social and behavioral issue—viz., how to identify and appropriately respond to concerns of possible child abuse.

Keywords: early childhood professional, professional development, online education, mandated reporter, experiential learning, child abuse

INTRODUCTION

Early childhood professionals (ECPs, a term used here to describe early childhood educators, daycare providers, childcare workers, preschool teachers, and others who work with young children) need better preparation to help them protect young children from abuse (Ayling et al., 2019; Dinehart & Kenny, 2015; McKee & Dillenburger, 2012). Only a very few trainings have any evidence-base for changing knowledge or attitudes (e.g., Gushwa et al., 2019; Kenny, 2007; Townsend & Haviland, 2016). For the vast majority, purported efficacy is based on self-report (e.g., “I feel like I learned a lot”), involves nonvalidated measures, and fails to evaluate for sustained effects on knowledge or attitudes, much less actual behavior (Ayling et al., 2019; Sokolowski, 2005). Even newer online training typically consists of little more than linear recitations of information regarding the various types of child abuse and their consequences, legal definitions and requirements for reporting, and contact information (Goldman & Evans, 2015; New York State Office of Child and Family Services,

n.d.; University of Pennsylvania, n.d.). It is in fact rather surprising that the potential of online learning has not been effectively leveraged for an issue as important and far reaching as child abuse.

Over 675,000 cases of child abuse (physical, sexual, and emotional abuse, and neglect) are confirmed annually in the U.S. (U.S. Department of Health & Human Services, 2018), with evidence that the true incidence is much higher (Finkelhor et al., 2010). The long-lasting and sometimes severe consequences of abuse include physical disabilities, cognitive impairment, neurological damage, mental health problems (depression, anxiety, posttraumatic stress, etc.), maladaptive behaviors (alcoholism, drug abuse, intimate partner violence), and of course further victimization (Norman et al., 2012; U.S. Department of Health & Human Services, 2018). In short, abuse can have a devastating impact on a child's life and the adult they become. Young children (0–5 years) are particularly vulnerable to victimization, accounting for more than 75% of deaths from abuse and a greater proportion of abuse than older

children for all categories of maltreatment except sexual abuse (U.S. Department of Health & Human Services, 2018). Despite 10 to 12 million American children being under their watchful eyes, ECPs identify fewer than 1% of all substantiated cases of child abuse in the United States each year (McKee & Dillenburger, 2012; U.S. Department of Health & Human Services, 2018).

There is little question that identifying and appropriately responding to suspected child abuse involves a number of complexities, but the dearth of well-grounded interventions to help people carry out this responsibility is a significant problem. This article describes how the online learning program, *iLookOut* for Child Abuse (*iLookOut*), was designed (conceptually and pedagogically) to help ECPs identify and appropriately respond to possible child abuse. Our hope is that sharing its structure and design can help others in higher education and learning technology develop evidence-based interventions that promote child well-being. *iLookOut* was developed under the aegis of Penn State's Center for the Protection of Children, with support from the National Institute of Child Health and Human Development. *iLookOut* is provided free of charge to ECPs and is currently available in the United States in the states of Pennsylvania and Maine.

EARLY CHILDHOOD PROFESSIONALS (ECPs)

The population of ECPs is diverse in terms of age, education, work setting, and available resources, but almost all are strategically positioned to identify and respond to child abuse. ECPs may be the only people outside of the immediate family to have extended opportunities to observe children on a daily basis, and thus they have the potential to both help prevent patterns of abuse from taking hold and act as key supports for children/families (Dinehart et al., 2013). Yet ECPs face considerable obstacles in recognizing signs of abuse, differentiating normal childhood injuries from abuse, and knowing when and how to report concerns about abuse. So it is not surprising that ECPs have identified “reporting possible abuse” as one of the most troubling ethical issue they face in their workplace (Ayling et al., 2019; Clyde & Rodd, 1989; Feeney & Sysko, 1986; Feng et al., 2009; McKenna, 2011).

Studies of ECPs' reporting habits for child abuse are relatively few compared with the many

studies involving other mandated reporters (e.g., teachers, nurses, doctors). The small body of research that examines the reporting experiences of ECPs (Bishop et al., 2002) reveals high levels of uncertainty about the decision to report, perceived “conflicts of loyalty” (Svensson & Janson, 2008), and a plethora of complexities that cause some ECPs to feel as if they are “dancing on the edge” (Feng et al., 2009). The challenges arise in part from ECPs' desire to preserve relationships with families and avoid causing harm while also meeting their legal, professional, and ethical responsibilities to protect children. One study found that ECPs are less likely to have ever reported child maltreatment compared to other professionals who work with children (Zellman & Bell, 1990), in part due to inadequate education regarding the circumstances and level of concern that warrant reporting (Ayling et al., 2019; Kenny, 2007). Even when ECPs do report, there is considerable report latency, (Svensson et al., 2015), with one older study finding an average lag time of 14 months between ECPs having suspicion and actually making a report (Sundell, 1997), a situation that if left unaddressed risks dire outcomes for many young children.

Professional training has been the chief mechanism for trying to improve ECPs' recognition and reporting of child abuse, the presumption being that increased knowledge will promote reporting. But due to the lack of rigorously evaluated interventions, little is known about the actual effect of education on ECPs' recognition of abuse, reporting behavior, or how to best prepare them to meet their responsibilities to protect children (Christian, 2008; Mathews et al., 2015). That said, research suggests that ECPs' lack of education contributes to their reports of suspected abuse having lower yields—with substantiation rates of just 6.3% compared to 25–33% for other mandated reporters (King et al., 2013; Pennsylvania Department of Human Services, 2017).

NEED FOR ONLINE INTERVENTIONS

Any intervention to help ECPs meet their professional, ethical, and legal responsibilities as mandated reporters must deal with multiple widely recognized challenges: wide variability in entry level training of ECPs; variability in the quality of professional development opportunities; logistical barriers to providing professional

development during working hours (be it a lack of down time during the workday or limited access to training); bureaucratic challenges to ensuring quality education across settings (from family-based daycares to corporate chains to church-based facilities); as well as short-staffing and 20%–40% annual turnover rates (Melusky et al., 2013; Whitebook et al., 2014). All these factors make it more difficult to establish childcare environments that are well prepared to protect young children (especially infants and toddlers) from harm (Zaslow, 2014).

The Advantages of Online Learning

That said, well-designed training programs can succeed, provided they deliver standardized, high quality curricula and reinforce learning (Lunenbergh, 2011). Online education has particular advantages for meeting the needs of ECPs and overcoming key challenges. Online learning lends itself to standardization is eminently scalable; provides ready, low-cost access to multimedia learning; can easily employ interactive exercises for experiential learning; can be accessed more flexibly than workshop based training; is as effective as in-person training at enhancing ECPs' knowledge, skills, and professional competencies; and provides ready means for tracking results (Ackerman, 2017; Barnes et al., 2018; Durden et al., 2016; Stone-MacDonald & Douglass, 2015).

To be effective, a training program needs to help ECPs understand the different forms of child abuse and their presenting signs and symptoms (Dinehart & Kenny, 2015), given that child abuse can be challenging to discern and perceptions of abuse can be mistaken (Christian & States, 2017; Reece & Christian, 2008). Effective training must also teach ECPs about their professional and legal responsibilities—which can be confusing (Mathews & Kenny, 2008; Mathews et al., 2015)—and also promote appropriate attitudes and behavior about reporting suspected abuse (Ajzen & Fishbein, 2005; Mathews et al., 2015). If done well, education can help ECPs be more aware, better prepared, and more inclined to appropriately recognize and effectively report suspected abuse (Crenshaw et al., 1995; Fraser et al., 2010; Yang et al., 2020). Importantly, this entails not only reducing missed cases of abuse but also minimizing over-reporting (Ho et al., 2017). Because child abuse is protean in its various presentations, it is at best misinformed

to suggest that ECPs should report whenever child abuse is “possible.” A child who is anxious or withdrawn or angry, or small, or tired, or slow to develop, or any number of other common things “may” be being abused. But it is neither warranted, practical, nor helpful to report every such child.

iLookOut for Child Abuse (iLookOut) Intervention

iLookOut is an online, interactive educational program that was designed to help better prepare ECPs for the challenging and important responsibility of protecting society's most vulnerable members from harm. A randomized controlled trial has shown that *iLookOut* improves ECPs' knowledge and changes ECPs' attitudes (in the desired direction) about child abuse and its reporting, and that ECPs very much like *iLookOut*. Previously published work describes those results and how the knowledge and attitudinal measures were developed and validated (Mathews et al., 2017), with more recent data demonstrating *iLookOut*'s efficacy in a real-world trial involving more than 11,000 ECPs (Yang et al., 2020). In the present article we describe the *iLookOut* learning program and its pedagogical underpinnings.

1. Understand and recognize possible child abuse
 - What does and does not constitute child abuse
 - How and where abuse occurs
 - Risk factors for abuse
 - Signs and symptoms of abuse
 - Consequences of abuse
 - Strategies and resources for responding to suspected abuse
2. Understand your responsibilities as a mandated reporters of suspected abuse
 - Legal responsibilities as a mandated reporter
 - Steps for carrying out these responsibilities
 - Consequences of failing to report suspected abuse
 - legal immunity so long as a report is made in good faith

Figure 1. *iLookOut*'s Didactic Learning Objectives

Learning Objectives

As with most standard training on reporting child abuse, *iLookOut*'s didactic learning objectives are to help ECPs:

1. Understand and recognize possible child abuse, and
2. Understand the responsibilities of being a mandated reporter (see Figure 1).

An additional learning objective, however, is affective in nature:

3. Become empowered and motivated to protect children who are at risk for abuse.

This includes helping ECPs engage with colleagues when concerns arise about the possibility of abuse, and helping them navigate barriers to reporting suspected abuse.

Structure of the Learning Program

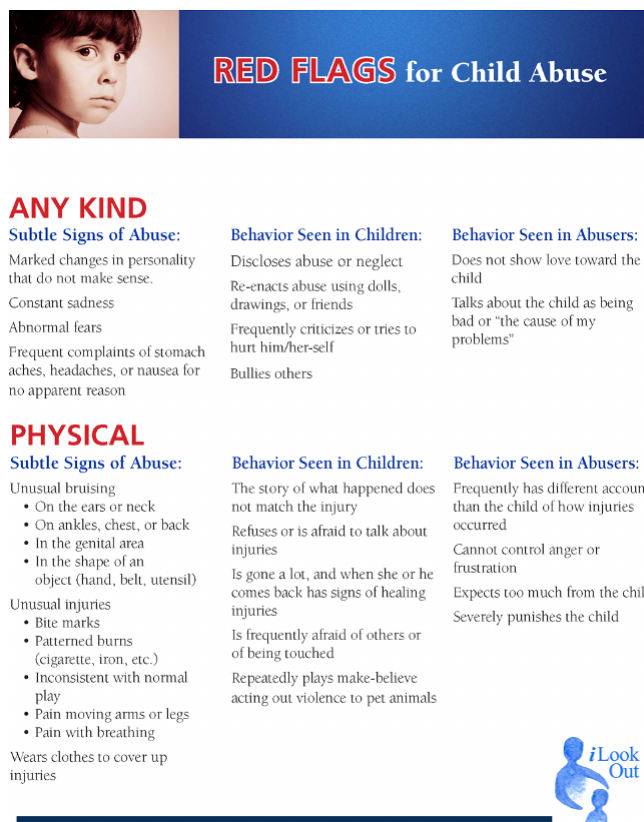
Learners first complete a registration section covering a broad range of demographic items (professional and personal), including prior experience with child abuse education and reporting suspected abuse. Learners then complete a pretest of knowledge (23 items) and attitudes (13 items) regarding child abuse and its reporting. As previously described (Mathews et al., 2017), the validated knowledge instrument examines whether learners know the signs and symptoms of abuse, the laws governing reporting, and the penalties for failing to report. The validated instrument assesses learners' views about reporting suspected abuse, including its value/utility, potential barriers, concern over legal liabilities and parental backlash, and their willingness to report over the objections of a supervisor. The Getting Started section begins with two separate videos of a young man and a young woman (both actors) talking about their experience of being abused as a child. While these testimonials do not contain any graphic descriptions or images of abuse, learners are alerted beforehand that "some people may find the videos upsetting, particularly if their own lives have somehow been affected by abuse." (Of note, no objections or negative comments about these testimonial videos have been received from any of the more than 13,000 ECPs who have completed *iLookOut*.) The purpose of these videos is to impress on learners the very real human impact that our actions (or failures to act) can have for children who are abused, as well as their families. At the end of the learning program, this connection is further reinforced with a brief video of a young girl who asks: "If you don't protect me, who will?"

Learner Interface

In this initial version of *iLookOut*, learners are asked to imagine being an early childhood educator working with four- and five-year-olds. Then, using an interactive, video-based storyline filmed with point-of-view videography (i.e., the camera functions as the learner's eyes), key events unfold through interactions with children, parents, and coworkers (all played by actors). As

more is revealed about each child's situation over the storyline's two-day time frame, the learner has to decide what, if any, action to take. In-depth information is provided for 5 of the children depicted, each of whom demonstrates risk factors for a particular form of abuse.

At different junctures in the story, resource files become available for learners to access attitudes. These include: 1) Types of Child Abuse (which defines and provides examples of each type of child abuse), 2) Red Flags for Abuse (see Figure 2), 3) Facts About Abuse (which includes state and national statistics about the incidence of abuse, physical and psychological consequences of abuse, tips for identifying abuse, legal penalties for failure to report suspected abuse, and links to useful websites), and 4) Reporting Suspected Abuse (which explains the threshold for reporting, how to operationalize the term "reasonable suspicion," and that mandated reporters have legal immunity so long as a report is made in good faith).



ANY KIND	PHYSICAL	Behavior Seen in Children:	Behavior Seen in Abusers:
Subtle Signs of Abuse: Marked changes in personality that do not make sense. Constant sadness Abnormal fears Frequent complaints of stomach aches, headaches, or nausea for no apparent reason	Subtle Signs of Abuse: Unusual bruising • On the ears or neck • On ankles, chest, or back • In the genital area • In the shape of an object (hand, belt, utensil) Unusual injuries • Bite marks • Patterned burns (cigarette, iron, etc.) • Inconsistent with normal play • Pain moving arms or legs • Pain with breathing Wears clothes to cover up injuries	Discloses abuse or neglect Re-enacts abuse using dolls, drawings, or friends Frequently criticizes or tries to hurt him/her-self Bullies others The story of what happened does not match the injury Refuses or is afraid to talk about injuries Is gone a lot, and when she or he comes back has signs of healing injuries Is frequently afraid of others or of being touched Repeatedly plays make-believe acting out violence to pet animals	Does not show love toward the child Talks about the child as being bad or "the cause of my problems" Frequently has different accounts than the child of how injuries occurred Cannot control anger or frustration Expects too much from the child Severely punishes the child

Figure 2. Red Flags for Abuse Handout

Learners also can choose to view additional videos and text files to learn more about the children portrayed in the storyline, including back stories

on the children and their families (see Figure 3). Part of the lesson here is that, as in real life, the more information one has, the better informed one's choices. That said, it is also made very clear that it is not the ECP's role to investigate whether child abuse has in fact occurred.



Figure 3. *iLookOut* Back-stories on Children

Throughout the *iLookOut* program, learners are posed questions, and based on their responses are provided information to augment (and/or correct) their understanding of child abuse and its reporting. At the end of the storyline, a video is shown in which narrators discuss elements within the storyline that could (or should) have raised or lowered the learner's concern about abuse for each of the children they encountered in the story. Learners are then asked to sign a pledge that they will fulfill their responsibilities as mandated reporters. They are also given follow-up materials (e.g., case scenarios for discussion, handouts) that they can print, download, and share with others in their work setting.

Learners then complete a knowledge and attitudes posttest (identical to the pretest) and a survey evaluating the learning experience. If any items on the knowledge test were answered incorrectly, the learner must identify the correct answer in order to complete the *iLookOut* program. The learning program then finishes with a disclaimer about the characters (children and adults) depicted in *iLookOut*, explaining that 1) *iLookOut* was written and casted to demonstrate diversity (race/ethnicity, gender, socio-economic status, awareness, temperament, etc.); 2) various of the depicted behaviors and personal characteristics may inadvertently reinforce negative stereotypes;

and 3) the creators welcome suggestions for avoiding such negative characterizations, provided they do not simply shift them to another group. To date, no complaints or suggestions on this matter have been received.

Completing *iLookOut* earns ECPs three hours of professional development credit and also satisfies their state requirement for mandated reporter training. ECPs' valuation of *iLookOut* is perhaps best demonstrated by the fact that more than 13,000 learners have completed it since open access began in January 2015, constituting a greater than 98% completion rate. Of these learners, more than 95% accessed all the supplemental resource files, and 94% reported being very satisfied with the learning program (mean = 8.8, where 10 = highest—manuscript forthcoming).

Pedagogical Approach

iLookOut was created by an interdisciplinary team with expertise in child abuse, pediatrics, early childhood education, online learning (including instructional designers and experts in teaching and learning technology), mandated reporter training, law, ethics, child advocacy, and victim services. Its interactive storyline was designed to 1) engage ECPs emotionally and intellectually, 2) increase their awareness about child abuse, and 3) help them feel both empowered and responsible to contact child protective services when there is reasonable suspicion of child abuse.

The opening testimonial videos set the stage for these goals. Prior to the first video, text appears on the screen rhetorically asking "Do you remember a time when YOU felt hurt and scared?... and nobody helped you?" Prior to the second video, the same question appears in text on the screen: "Do you remember a time when YOU felt hurt and scared?" But this time, the follow-on text reads: "... and someone asked the right questions and did the right thing to protect you?" The purpose of this sequence is to encourage learners not only to identify with the vulnerability of children who are at-risk but also to identify as being part of a system that, however imperfect, is the only system we have for protecting children. As the learning program then moves into the video-based storyline, ECPs are immersed in realistic scenarios that both teach them new information and challenge them to put into practice what they are learning. In addition to helping learners operationalize new information,

this approach encourages important affective skills to encourage ECPs to be more proactive in protecting real children from harm.

This is one example of how *iLookOut* is grounded in an experiential learning, conceptual model that is a key feature of adult learning theory. Drawing on the work of Knowles (1984), Billington (2000), and Kolb (2009), *iLookout's* design recognizes that adults learn best when 1) they know why they need to learn the material, 2) the learning process is experiential, 3) learning is framed as problem-solving, and 4) the material to be learned has immediate value (Knowles, 1984). In keeping with Billington's key factors for promoting adult development, *iLookOut* 1) challenges ECPs just beyond their present level of ability so they are pushed to grow but not pushed so far that they give up; 2) uses exercises to reinforce facts and frameworks regarding suspected child abuse; and 3) allows learners to proceed and digest information at their own pace. Because *iLookOut* is accessible 24/7 and can be paused/resumed as desired, including across multiple sessions, it also leverages ECPs' preference for flexibility in professional development (Kyzar et al., 2014). Additionally, *iLookOut's* online platform provides an emotionally safe environment for experiential learning, which has been shown to improve knowledge acquisition and implementation among ECPs (Kyzar et al., 2014).

By interweaving an interactive storyline with didactic information, decision-points, and critical feedback to learners' responses, *iLookOut* aims to reflect adult learning best practices (Billington, 2000) and embody the key elements of Kolb's experiential model (Kolb & Kolb, 2009). In the context of helping ECPs become responsible, mandated reporters of child abuse, these key elements manifest as follows:

- **“Concrete Experience”**—helping ECPs reinterpret experiences they have previously encountered.
- **“Reflective Observation”**—helping ECPs consider and problem-solve any tension/conflict between the lived experience of being a ECP and their responsibilities to protect children.
- **“Abstract Conceptualization”**—promoting reflection about the meaning/implications of concepts such as abuse and suspicion.

- **“Active Experimentation”**—providing practice opportunities to apply new information/understanding.

Taken together, these elements are intended to provide a nuanced, yet practical, educational experience expressly designed for early childhood educators and caregivers. In particular, *iLookout's* interactive storyline, pairing of questions with immediate critical feedback, posttest reinforcement, reflective debriefing, and follow-up engagement capitalize on Kolb's observation that critical thinking skills develop best when learners transform their own experience into knowledge by acting on their learning.

iLookOut's Evidence Base

The *iLookOut* program is integrated with a learning management system that tracks pre- and posttest data, responses to questions within the learning program and a postprogram evaluation of the learning experience. In both randomized controlled trials (Mathews et al., 2017; National Library of Medicine, n.d.) and a real-world study (Yang et al., 2020), *iLookOut* has been shown to significantly improve knowledge (effect size = 0.96) and change attitudes (effect size = 0.8), such that learners are more predisposed to seeing reporting as the right thing to do when they suspect that a child may have been abused. In an ongoing randomized controlled trial (Humphrey et al., 2021), the interactive, experiential learning approach employed by *iLookOut* resulted in significantly higher effect sizes regarding both knowledge (1.09) and attitudes (0.67) compared to standard didactic online training (0.67 and 0.54, respectively). Additionally, preliminary data suggest that reports from ECPs who complete *iLookOut* are more likely to be screened-in for further consideration, and when formally assessed they are more likely to result in findings of abuse and/or social services being recommend for the child/family in question. Because the learner's responses within the various learning modules are also recorded by the learning management system, it is also possible to study patterns of learning; however, this has not been a focus of inquiry to date.

By design, all legal/policy-related content in *iLookOut* is contained within discreet learning modules (as opposed to the video-based storyline). So, while *iLookOut* was originally developed

for use in Pennsylvania, state-specific content can be readily revised to comport with legal and policy-related requirements of other states, as has already been done for the state of Maine. This adaptability provides opportunities to study the efficacy of *iLookOut* with diverse populations and to experiment with different ways of framing the learning material.

IMPLICATIONS FOR PRACTICE

Where child well-being is concerned, educational interventions should aim not only to impart knowledge but also to help learners feel more empowered and motivated to act when needed. Because interactive, story-based learning can help engage learners and facilitate skill-building, it is a particularly promising approach for achieving such affective and behavioral goals. Moreover, when experiential learning is not only problem-based but practical in its application, learners are more likely to appreciate the relevance of what is being taught and draw the connections needed to apply it in real life. Online interventions have the additional advantage of being accessible, adaptable, and open to analysis, as well as a ready mechanism for promoting a greater, shared understanding about how to effectively support child well-being.

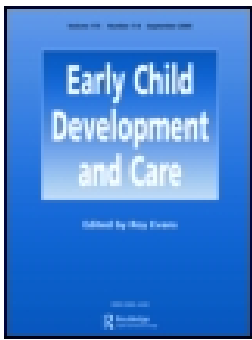
For projects with even quite modest budgets there are several general techniques for optimizing learner engagement and increasing scholarship on which elements are best suited to various purposes (Boller & Kapp, 2017). Such techniques include introducing characters and/or storylines that are likely to resonate with the perspectives and experiences of learners, providing (safe) opportunities for learners to practice and fail at applying newly acquired knowledge, and employing digital badging as a way to leverage people's desire for recognition. Advances in technology have also made immersive educational experiences increasingly affordable through the use of off-the-shelf animated scenarios that can be tailored to the desired setting and content such as Gamelearn (www.game-learn.com). Such gamification techniques can be very effective at augmenting educational interventions, particularly when coupled with a clear understanding of adult learning theory.

REFERENCES

- Ackerman, D. J. (2017). Online child care training in the United States: A preliminary investigation of who participates, what is offered, and on which topics the workforce is focusing. *International Journal of Child Care and Education Policy*, 11(1), article 12. <https://doi.org/10.1186/s40723-017-0037-7>
- Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behaviour. In D. Albarracin, B. Johnson, & M. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Lawrence Erlbaum Associates Publishers.
- Ayling, N. J., Walsh, K., & Williams, K. E. (2019). Factors influencing early childhood education and care educators' reporting of child abuse and neglect. *Australasian Journal of Early Childhood*, 45(1), 95–108. <https://doi.org/10.1177/1836939119885307>
- Barnes, J. K., Guin, A., & Allen, K. (2018). Training needs and online learning preferences of early childhood professionals. *Journal of Early Childhood Teacher Education*, 39(2), 114–130. <https://doi.org/10.1080/10901027.2017.1347909>
- Billington, D. D. (2000). Seven characteristics of highly effective adult learning programs. *New Horizons for Learning*. <http://education.jhu.edu/PD/newhorizons/lifelonglearning/workplace/articles/characteristics>
- Bishop, A., Lunn, P., & Johnson, K. (2002). 'I would just like to run away and hide, but I won't!' Exploring attitudes and perceptions on child protection issues with early years teacher trainees on the threshold of their careers. *Westminster Studies in Education*, 25(2), 187–199. <https://doi.org/10.1080/0140672020250208>
- Boller, S., & Kapp, K. (2017). Play to learn: Everything you need to know about designing effective learning games. Association for Talent Development.
- Christian, C. W. (2008). Professional education in child abuse and neglect. *Pediatrics*, 122, S13–S17. <https://doi.org/10.1542/peds.2008-0715f>
- Christian, C. W., & States, L. J. (2017). Medical mimics of child abuse. *American Journal of Roentgenology*, 208(5), 982–990.
- Clyde, M., & Rodd, J. (1989). Professional ethics: There's more to it than meets the eye! *Early Child Development and Care*, 53(1), 1–12. <https://doi.org/10.1080/0300443890530101>
- Crenshaw, W. B., Crenshaw, L. M., & Lichtenberg, J. W. (1995). When educators confront child abuse: An analysis of the decision to report. *Child Abuse & Neglect*, 19(9), 1095–1113. [https://doi.org/10.1016/0145-2134\(95\)00071-F](https://doi.org/10.1016/0145-2134(95)00071-F)
- Dinehart, L. H., Katz, L. F., Manfra, L., & Ullery, M. A. (2013). Providing quality early care and education to young children who experience maltreatment. *Early Childhood Education Journal*, 41(4), 283–290. <https://doi.org/10.1007/s10643-012-0553-6>
- Dinehart, L., & Kenny, M. C. (2015). Knowledge of child abuse and reporting practices among early care and education providers. *Journal of Research in Childhood Education*, 29(4), 429–443. <https://doi.org/10.1080/02568543.2015.1073818>
- Durden, T., Mincemoyer, C., Crandall, L., Alviz, K., & Garcia, A. (2016). Gateway to quality: Online professional development for family childcare providers. *Early Child Development and Care*, 186(7), 1079–1095. <https://doi.org/10.1080/03004430.2015.1076400>
- Feeney, S., & Sysko, L. (1986). Professional ethics in early childhood education: Survey results. *Young Children*, 42(1), 15–22.
- Feng, J.-Y., Chen, S.-J., Wilk, N. C., Yang, W.-P., & Fetzer, S. (2009). Kindergarten teachers' experience of reporting child abuse in Taiwan: Dancing on the edge. *Children and Youth Services Review*, 31(3), 405–409. <https://doi.org/10.1016/j.childyouth.2008.09.007>
- Finkelhor, D., Turner, H., Ormrod, R., & Hamby, S. L. (2010). Trends in childhood violence and abuse exposure. *Archives of Pediatric Adolescence Medicine*, 164(3), 238–242. <https://doi.org/10.1001/archpediatrics.2009.283>
- Fraser, J. A., Mathews, B., Walsh, K., Chen, L., & Dunne, M. (2010). Factors influencing child abuse and neglect Recognition and reporting by nurses: A multivariate analysis. *International Journal of Nursing Studies*, 47(2), 146–153. <https://doi.org/10.1016/j.ijnurstu.2009.05.015>
- Goldman, S., & Evans, N. (2015). Child abuse mandated reporter training for New York State. *Wild Iris Medical Education, Inc.*
- Gushwa, M., Bernier, J., & Robinson, D. (2019) Advancing child sexual abuse prevention in schools: An exploration of the effectiveness of the Enough! online training program for K–12 Teachers. *Journal of Child Sexual Abuse*, 28(2), 144–159. <https://doi.org/10.1080/10538712.2018.1477000>
- Ho, G. W. K., Gross, D. A., & Bettencourt, A. (2017). Universal mandatory reporting policies and the odds of identifying child physical abuse. *American Journal of Public Health*, 107(5), 709–716. <https://doi.org/10.2105/AJPH.2017.303667>
- Humphreys, K. L., Piersiak, H. A., Panlilio, C. C., Lehman, E. B., Verdiglione, N., Dore, S., & Levi, B. H. (2021). A randomized control trial of a child abuse mandated reporter training: Knowledge and attitudes. *Child Abuse & Neglect*, 117, <https://doi.org/10.1016/j.chiabu.2021.105033>.
- Kenny, M. C. (2007). Web-based training in child maltreatment for future mandated reporters. *Child Abuse and Neglect*, 31(6), 671–678. <https://doi.org/10.1016/j.chiabu.2006.12.008>
- King, B., Lawson, J., & Putnam-Hornstein, E. (2013). Examining the evidence: Reporter identity, allegation type, and sociodemographic characteristics as predictors of maltreatment substantiation. *Child Maltreatment*, 18(4), 232–244. <https://doi.org/10.1177/1077559513508001>

- Knowles, M. (1984). *The adult learner: A neglected species* (3rd ed.). Gulf Publishing
- Kolb, A. Y., & Kolb, D. A. (2009). The learning way: Meta-cognitive aspects of experiential learning. *Simulation and Gaming*, 40(3), 297–327. <https://doi.org/10.1177/1046878108325713>
- Kyzar, K. B., Chiu, C., Kemp, P., Aldersey, H. M., Turnbull, A. P., & Lindeman, D. P. (2014). Feasibility of an online professional development program for early intervention practitioners. *Infants and Young Children*, 27(2), 174–191. <https://doi.org/10.1097/IYC.0000000000000007>
- Lunenburg, F. C. (2011). Curriculum models for preschool education: Theories and approaches to learning in the early years. *Schooling*, 2(1), 1–6.
- Mathews, B., & Kenny, M. (2008). Mandatory reporting legislation in the United States, Canada and Australia: A cross-jurisdictional review of key features, differences and issues. *Child Maltreatment*, 13(1), 50–63. <https://doi.org/10.1177/1077559507310613>
- Mathews, B., Walsh, K., Coe, S., Kenny, M. C., & Vagenas, D. (2015). Child protection training for professionals to improve reporting of child abuse and neglect [intervention protocol]. *Cochrane Database of Systematic Reviews*, (6), Art. No.: CD011775. <https://doi.org/10.1002/14651858.CD011775>
- Mathews, B., Yang, C., Lehman, E. B., Mincemoyer, C., Verdiglione, N., & Levi, B. H. (2017). Educating early childhood care and education providers to improve knowledge and attitudes about reporting child maltreatment: A randomized controlled trial. *PLoS One*, 12(5). <https://doi.org/10.1371/journal.pone.0177777>
- McKee, B. E., & Dillenburger, K. (2012). Effectiveness of child protection training for pre-service early childhood educators. *International Journal of Educational Research*, 53, 348–359. <https://doi.org/10.1016/j.ijer.2012.04.008>
- McKenna, A. (2011). *Reluctant to report: The mandated reporting practices of child care providers* (Doctoral dissertation, Western Michigan University). ScholarWorks at WMU. Dissertations. 437. <https://scholarworks.wmich.edu/dissertations/437>
- Melusky, R., Slavinsky, E., Strauser, M., & Le, A. (2013). *The 50-state child care licensing study 2011–2013*. Lexington (KY): National Association for Regulatory Administration. http://www.naralicensing.org/assets/docs/ChildCareLicensingStudies/2011-2013_child%20care%20licensing%20study.pdf
- National Library of Medicine. (n.d.). iLook Out for Child Abuse: An Innovative Learning Module for Childcare Providers. ClinicalTrials.gov, NCT03185728. Retrieved from <https://clinicaltrials.gov/ct2/results?cond=&term=NCT03185728&cntry=&state=&city=&dist=>
- New York State Office of Child and Family Services (n.d.). Mandated Reporter Resource Center. <https://nysmandatedreporter.org/TrainingCourses.aspx>.
- Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse. Emotional Abuse, and Neglect. *PLoS Medicine*, 9(11). <https://doi.org/10.1371/journal.pmed.1001349>
- Pennsylvania Department of Human Services (2017). *Child Protective Services 2017 Annual Report*. Available at <https://www.dhs.pa.gov/docs/Publications/Pages/Child-Abuse-Reports.aspx>
- Reece, R. M., & Christian, C. (2008). *Child abuse: Medical diagnosis and management* (3rd ed.). American Academy of Pediatrics.
- Sokolowski, S. W. (2005). Pennsylvania Family Support Alliance, Mandated Reporter Training—Program Evaluation, Summary Report of Findings. <https://app.pa-fsa.org/Portals/0/Docs/Mandated-Reporters/Report%20of%20Findings%20Mandated%20Reporter%20Training.pdf>
- Stone-MacDonald, A., & Douglass, A. (2015). Introducing online training in an early childhood professional development system: Lessons learned in one state. *Early Childhood Education Journal*, 43, 241–248. <https://doi.org/10.1007/s10643-014-0649-2>
- Sundell, K. (1997). Child-care personnel's failure to report child maltreatment: Some Swedish evidence. *Child Abuse and Neglect*, 21(1), 93–105. [https://doi.org/10.1016/S0145-2134\(96\)00133-0](https://doi.org/10.1016/S0145-2134(96)00133-0)
- Svensson, B., & Janson, S. (2008). Suspected child maltreatment: Preschool staff in a conflict of loyalty. *Early Childhood Education Journal*, 36(1), 25–31. <https://doi.org/10.1007/s10643-008-0248-1>
- Svensson, B., Andershed, H., & Janson, S. (2015). A survey of Swedish teachers' concerns for preschool children at risk of maltreatment. *Early Childhood Education Journal*, 43(6), 495–503. <https://doi.org/10.1007/s10643-014-0684-z>
- Townsend, C., & Haviland, M. (2016). The impact of child sexual abuse training for educators on reporting and victim outcomes: The Texas Initiative. Charleston, S.C., Darkness to Light. <https://www.d2l.org/wp-content/uploads/2017/03/The-Impact-of-CSA-Training-for-Educators.pdf>
- University of Pennsylvania. (n.d.). Recognizing and Reporting Child Abuse: Mandated and Permissive Reporting in Pennsylvania Online Training. <https://www.hr.pitt.edu/clearances-training/training>

- U.S. Department of Health & Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. (2018). Child Maltreatment 2016. Available from <https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/child-maltreatment>
- Whitebook, M., Phillips, D., & Howes, C. (2014). Worthy work, STILL unlivable wages: The early childhood workforce 25 years after the National Child Care Staffing Study. Center for the Study of Child Care Employment, University of California, Berkeley. <https://cscce.berkeley.edu/worthy-work-still-unlivable-wages/>
- Yang, C., Panlilio, C., Verdiglione, N., Lehman, E. B., Hamm, R. M., Fiene, R., Dore, S., Bard, D. E., Grable, B., & Levi, B. (2020). Generalizing findings from a randomized controlled trial to a real-world study of the iLookOut, an online education program to improve early childhood care and education providers' knowledge and attitudes about reporting child maltreatment. *PloS One*, 15(1), e0227398. <https://doi.org/10.1371/journal.pone.0227398>
- Zaslow, M. (2014). General features of effective professional development. In H. P. Ginsburg, M. Hyson, & T. A. Woods; (Eds.), *Preparing early childhood educators to teach math: Professional development that works* (pp. 97–115). Brookes Publishing Co.
- Zellman, G. L., & Bell, R. M. (1990). The role of professional background, case characteristics, and protective agency response in mandated child abuse reporting. RAND Corp.



Ilookout for child abuse: conceptual and practical considerations in creating an online learning programme to engage learners and promote behaviour change

Benjamin H. Levi, Andrew Belser, Karl Kapp, Nicole Verdiglione, Claudia Mincemoyer, Sarah Dore, Jane Keat & Richard Fiene

To cite this article: Benjamin H. Levi, Andrew Belser, Karl Kapp, Nicole Verdiglione, Claudia Mincemoyer, Sarah Dore, Jane Keat & Richard Fiene (2019): Ilookout for child abuse: conceptual and practical considerations in creating an online learning programme to engage learners and promote behaviour change, Early Child Development and Care, DOI: [10.1080/03004430.2019.1626374](https://doi.org/10.1080/03004430.2019.1626374)

To link to this article: <https://doi.org/10.1080/03004430.2019.1626374>



Published online: 07 Jun 2019.



Submit your article to this journal [↗](#)



View Crossmark data [↗](#)



llookout for child abuse: conceptual and practical considerations in creating an online learning programme to engage learners and promote behaviour change

Benjamin H. Levi ^a, Andrew Belser^b, Karl Kapp^c, Nicole Verdiglione^d, Claudia Mincemoyer^e, Sarah Dore^f, Jane Keat ^g and Richard Fiene^{h,i}

^aDepartments of Humanities & Pediatrics, Penn State College of Medicine & Penn State Children's Hospital, Hershey, PA, USA; ^bTheatre Department, Penn State University, State College, PA, USA; ^cInstructional Technology Department, Bloomsburg University, Bloomsburg, PA, USA; ^dDepartment of Humanities, Penn State College of Medicine, Hershey, PA, USA; ^eExtension Education Department, Penn State University, State College, PA, USA; ^fDepartment of Humanities, Penn State College of Medicine, Hershey, PA, USA; ^gEducation Department, Penn State University, Middletown, PA, USA; ^hDepartments of Psychology & Human Development, Penn State University, Elizabethtown, PA, USA; ⁱResearch Institute for Key Indicators, Elizabethtown, Pennsylvania

ABSTRACT

iLookOut for Child Abuse (iLookOut) is an online, interactive educational programme designed to help early childhood educators/practitioners protect young children from harm, and in particular to become better at identifying and reporting suspected child abuse. The purpose of this paper is to describe the conceptual foundations and practical considerations that guided the development of iLookOut.

ARTICLE HISTORY

Received 11 May 2019

Accepted 29 May 2019

KEYWORDS

Online learning; child abuse; gamification; early childhood; mandated reporting; micro-learning

Introduction

We know that young children (0–5 years) are more vulnerable to victimization. They account for >75% of deaths from abuse, and a greater proportion of abuse than older children for all categories of maltreatment except sexual abuse (Administration on Children, Youth and Families, 2016). Because early childhood professionals (ECPs)¹ have extended opportunities to observe children and get to know their families, ECPs are strategically positioned to both act as key supports for children/families, and help prevent patterns of maltreatment from taking hold – which includes reporting suspected abuse (Dinehart, Katz, Manfra, & Ullery, 2013). Yet despite 10–12 million American children being under their care, ECPs identify <1% of the more than 650,000 substantiated cases of child abuse in the United States each year (Administration on Children, Youth and Families, 2016).

The earlier in life child abuse occurs, and the longer it persists, the more devastating and lasting its consequences are likely to be for children and the adults they become (Norman et al., 2012). Effective intervention requires knowing which children are at risk – which in turn requires awareness, knowledge of what to report, and appropriate action. Yet, simply presenting information to people who care for children is unlikely to result in significant changes in their behaviour. Emily Dickinson provided both rhyme and reason for taking a more nuanced approach when she commended us to 'Tell all the truth, but tell it slant ... Or every man be blind' (Franklin, 1999). Dickinson's insight is that to bring about the change we envision, we must find creative ways to engage others to do the right thing. As this applies to child abuse, it is not enough to educate those entrusted with

protecting children; we must do so in a way that both prepares them to exercise good judgment, and motivates them to act accordingly.

To engage ECPs, *iLookOut* uses a video-based storyline and game-based techniques (in conjunction with pre/post-testing and follow-up activities). Because *iLookOut* is online, it provides ready, low-cost access to multi-media learning 24/7, and can be paused/resumed as desired (including across multiple sessions). *iLookOut*'s format (which includes an overall 8th grade reading level) was designed to provide standardized, high quality education to a ECP workforce that is known for wide variability in its entry level training, skill-sets, work setting, and opportunities for professional development (National Association for Regulatory Administration, 2013; Whitebook, Phillips, & Howes, 2014). To date, *iLookOut* has been shown to significantly improve ECPs' knowledge and change their attitudes (in the desired direction) about child abuse and its reporting (Mathews et al., 2017; Yang et al., 2019); and preliminary results from a 5-year randomized controlled trial suggest that *iLookOut* has a positive impact on ECPs' actual reporting of suspected abuse (see ClinicalTrials.gov, NCT03185728). The purpose of this paper is to describe the conceptual foundations and practical considerations that guided the development of *iLookOut*.

Early stages of *iLookOut*

A pilot version of *iLookOut* was designed by a multi-disciplinary team (working mostly pro-bono) that included experts in mandated reporter training, child advocacy, law, child abuse, pediatrics, early childhood education, ethics, instructional design, and online learning. The initial storyline used point-of-view videography (i.e. the camera functioning as the learner's eyes), with learners taking the role of a ECP working with 4 and 5 year-olds. Key events unfolded over the storyline's 2-day time-frame, portrayed through video-based interactions with children, parents, and co-workers (all played by non-professional actors). As more was revealed about each child's situation, the learner had to decide what, if any, action to take. At different junctures in the story, resource files (e.g. *Types of Child Abuse*; *Facts About Abuse*; *Red Flags for Abuse*; *Reporting Suspected Abuse*; etc.) became available for learners to read, print, and/or download for future reference. Learners were also posed questions and then, based on their responses, were provided information to correct or augment their existing understanding of child abuse and its reporting. At the end, learners were shown a video with narrators discussing elements within the storyline that could (or should) have altered the learner's concern about abuse for each of the 5 children highlighted in the story. Learners were then asked to sign a pledge to fulfil their responsibilities as mandated reporters, and given follow-up resources (e.g. case scenarios for discussion, handouts, etc.) for themselves or co-workers to use.

As a prototype, this initial version of *iLookOut* (which can be accessed at <https://www.ilookoutforchildabuse.com/>) served its purpose and was well embraced – with >12,000 Pennsylvania ECPs completing *iLookOut* in the 3 years since open access began. But with external funding (1 R01 HD088448-03) it became possible to introduce additional creative elements to more effectively engage learners (emotionally and intellectually), develop their skills and judgment, and (hopefully) motivate and empower them to do the right thing. To bring this about, we brought in experts in theatre and film, gamification, ECP professional development, and decision-making psychology.

Conceptual and practical enhancements

Storyline

The first major enhancement involved framing the storyline from different points of view. So rather than positioning the learner as the protagonist who is under pressure to make the right decision at each step, the learner enters at a slant. This involved adapting the script by adding two characters – Megan, a young ECP grappling with some concerns about a child she takes care of, and Elisha, an 'expert' ECP and former mentor of Megan who has experience navigating such situations. In the

introductory video, Elisha, who has just gotten off the phone with Megan, introduces herself and invites the learner to join her to meet with Megan at a nearby coffee shop (see [Figure 1](#)). By creating this personal relationship with Elisha, the learner is drawn into Megan's concerns (about a child who is potentially being abused) without being pressured to make the right decision for an at-risk child. Megan's story and concerns are then revealed through conversation with Elisha, which at times involves video-flashbacks of events at Megan's childcare facility. Integrated throughout Megan and Elisha's conversation are 'lessons' that Elisha shares with the learner, sometimes as an aside, other-times coinciding with advice and instruction given to Megan. The learning programme is tied together through the story that unfolds during conversation at the coffee shop. Because the learner is part of this conversation, Megan and Elisha's real-world concerns become the learner's concerns as well – with Megan's emotional responses guiding the learner through the experience. This includes relating to Megan's fears about how (as 'just a ECP') she can know (and do) what is needed to protect a child from abuse. Elisha's skilful mentoring helps learners become more knowledgeable about red flags for abuse, as well as when and how to report their concerns ([Figure 2](#)). Moreover, the interaction models how learners can carry out their professional responsibilities when they are concerned (and/or uncertain) about a child's safety.

Prior to filming the video-sequences, the *iLookOut* team spent considerable time casting and then rehearsing professional actors to help them establish a strong connection to their character's role; develop a clear understanding of the story's purpose as it concerned ECPs; and deliver nuanced portrayals of characters facing different sides of an abuse situation.

Game elements and techniques

An emerging instructional innovation to enhance learning and motivation is the use of game elements, game mechanics, and game-based thinking (Boller & Kapp, 2017). Gamification is effective because the 'underlying dynamics that make games engaging are largely already recognized and utilized in modern pedagogical practices, [albeit] under different designations' (Stott &

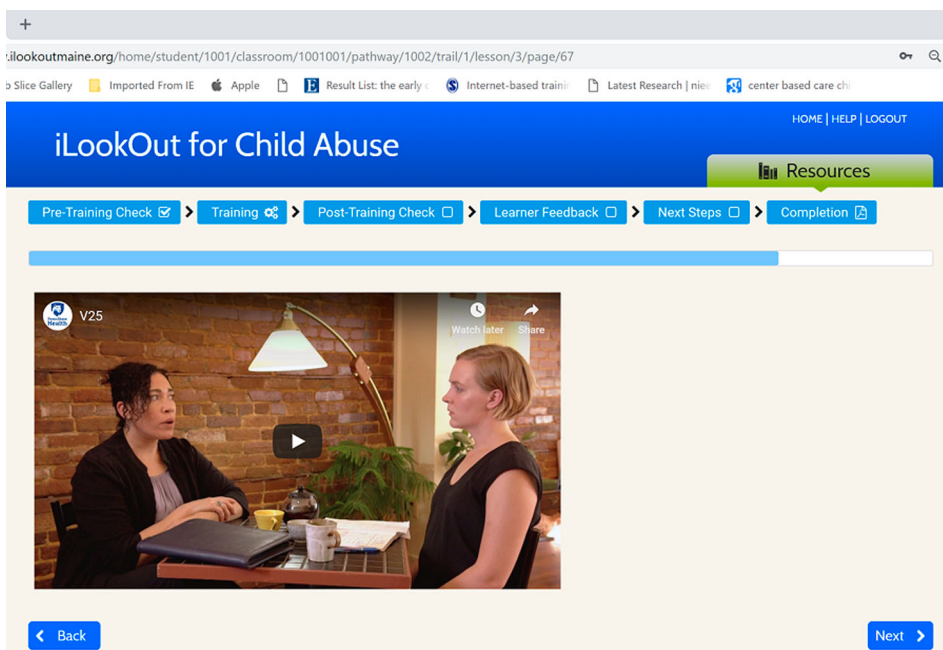


Figure 1. Interaction between Elisha a younger early childhood professional, Megan.



Figure 2. Red Flags handout.

Neustaedter, 2013). As such, the following game-elements were chosen to motivate ECPs to participate in *iLookOut*, help them retain what they learned, and apply this learning in real-world settings.

Story

Story constructs are commonly used in games to add meaning, provide context, and guide action (Kapp, 2012; Martin & Powers, 1983). Stories evoke emotions and provide a context for placing information; they are how humans have handed down information for millennia. People frequently use narratives to inform views of themselves and others, frame past experiences, and provide meaning to actions and activities (McAdams, 2006).

Characters

The use of characters is an important element in many games, especially those incorporating video game dynamics – where characters are often used to motivate players to engage more and/or try harder to succeed. A general principle of learning games is that the more a learner can identify with a character's situation and experience, the more likely s/he is to become invested in the learning exercise.

Mystery

Introducing informational complexity and novel, exciting situations can create a sense of mystery, stimulating learners' curiosity to uncover important information (Malone, 1981; Wilson et al., 2009). By unveiling information gradually, gamification techniques encourage individuals to stay engaged in learning activities, and prevent 'too much' information being revealed at once – which can impede learning. *iLookOut*'s evolving storyline stimulates learners, and also provides a staged sequence for developing their awareness, knowledge base, and motivation to act.

Badging

Badges are visible signs of having accomplished particular tasks or performed various behaviours. Badges in gamification provide positive reinforcement to learners for such accomplishments, along with visible credentialing that a learning objective has been achieved (Gibson, Ostashewski,

Flintoff, Grant, & Knight, 2015). With *iLookOut*, ECPs receive digital badges upon completing the initial 3-hour learning programme, as well as for follow-up advanced training that involves micro-learning activities (see below).

Content

The content of *iLookOut* was enhanced to include several topics that are often not explored in depth, despite their importance for understanding when and how to report suspected child abuse. This was done by creating discrete learning modules (within the learning programme) whose interactive exercises use game elements and mechanics. For example, one learning module integrates previous scholarship on 'reasonable suspicion' (Levi & Loeben, 2004; Levi & Portwood, 2011; Levi, Dellasega, & Boehmer, 2012) to construct an animated algorithm that provides both conceptual depth and guidance for figuring out whether a given situation warrants reporting. A separate learning module helps learners differentiate 'child abuse' from parenting that is 'less than ideal.' Other learning modules explore factors (e.g. intimate partner violence and animal abuse) that should raise concern for child abuse. Other modules help differentiate birthmarks from bruising, and worrisome child behaviours from normal development. One of the more lengthy learning modules helps teach ECPs about when and how to ask probing questions (see below). Additionally, several modules explore the mechanics of reporting – what information to have at the ready, how to articulate one's concerns, and what a phone call to child protective services will likely entail (see Figure 3). Throughout these various learning modules, best practices are reinforced, such as remaining alert for patterns of injury or unusual behaviour, and documenting one's observations.

Generalizability

iLookOut is designed for early childhood professionals, and is currently being deployed in the state of Maine (with an earlier prototype still in use in Pennsylvania). Because script-writing, acting, filming,

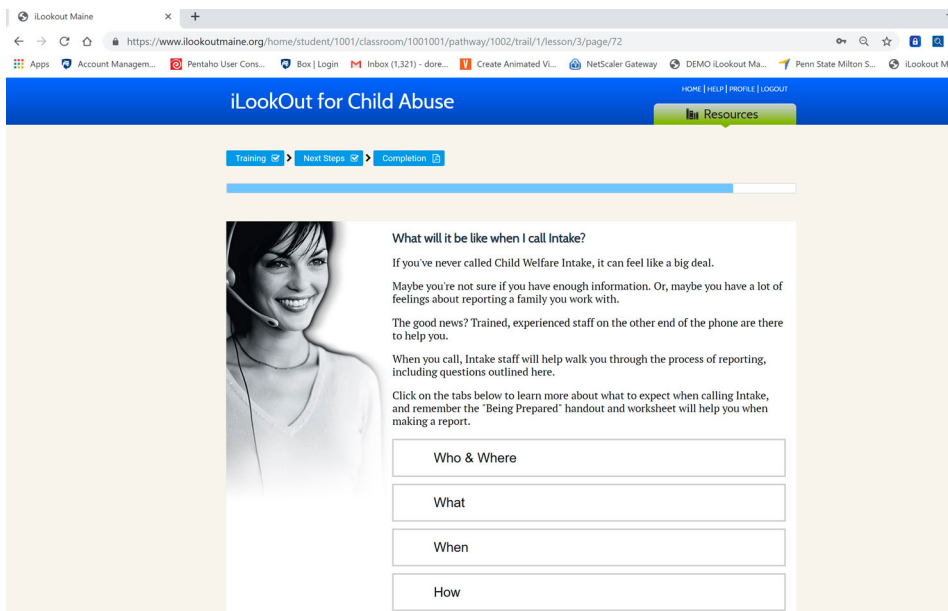


Figure 3. Interactive exercise to help learners know what to expect when calling child protective services.

and editing/post-production are time-consuming and expensive, the video-based storyline for *iLookOut* was constructed to be compatible with all U.S. states, and to be applicable to a broad scope of ECPs. Specifically, the setting and events for the storyline were chosen to reflect the lived experience of ECPs who work in any of a variety of early childhood programmes – family-based, centres, pre-schools, and Head Start programmes. In order for the terminology in *iLookOut* to have the broadest applicability, the study team reviewed key U.S. Children's Bureau white papers, and examined (from all 50 states) child welfare guidelines, and relevant policies regarding early childhood professionals. These findings were then used to construct *iLookOut's* storyline/script so it would be appropriate for a wide array of ECPs throughout the U.S.

By contrast, the more specific 'learning modules' were constructed to be readily adaptable to state-based differences. So, while large swaths of content within the various learning modules are *not* state-specific, some aspects are. For example, definitions, penalties, reporting requirements, contact information, web-based resources, and so forth vary from state to state. For this reason, the discrete learning modules were designed so they could be readily revised to accord with the specific laws, policies, and informational needs of any given U.S. state.

The pre/post-knowledge test was likewise constructed to accommodate state-based differences (Yang et al., 2019). As described elsewhere, a multi-disciplinary group of experts identified key precepts and formed them into question-items, which were then vetted by a separate group of content experts, and subsequently subjected to a formal think-aloud validation protocol (Panlilio et al., 2019). This iterative process yielded a knowledge test that is readily adaptable for use in any U.S. state.

Cognitive sequencing map

To track all curriculum content and ensure integration of learning objectives, a 'cognitive mapping' process was undertaken (Kapp et al., 2019). Cognitive Mapping has been found to be useful in many different arenas (e.g. health planning research, Stadler, Dugmore, Venables, MacPhail, & Delany-Moretwe, 2013, and engineering, Dixon & Lammi, 2014) to 'map' the cognitive territory being covered. Cognitive maps can demonstrate the linear progression of a concept or project, or the non-linear relationships between various factors. But they also can serve a developmental purpose, helping people (be they researchers, policy makers, teachers, or learners) develop a deeper understanding of how different elements within a given programme are (or should be) related to one another.

Cognitive mapping of the *iLookOut* programme was used, first and foremost, to identify and ensure the integration of its educational content, skill-building exercises, and metrics for learner achievement. Additionally, *iLookOut's* cognitive map catalogues the various teaching techniques used in different learning modules, and also helps elucidate elements that were initially missing – be it specific topics (e.g. the relationship between substance abuse disorders and child abuse); conceptual matters (e.g. techniques for avoiding emotional triggering); or opportunities to expand on existing content (e.g. how to promote resilience, or support families where intimate partner violence is occurring). Finally, cognitive sequencing provides a prototype for how ECPs will, themselves, sequence and connect *iLookOut's* content and learning goals.

Question asking

We know that experts approach situations differently from novices (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Ericsson & Pool, 2016; Kalyuga & Sweller, 2018). In particular, experts typically consider and explore different kinds of 'questions,' which allows them to both learn information and discover connections that novices do not. One of our central goals was to help ECPs develop their ability to ask better questions, and better understand when and to whom various questions should (or should not) be asked. Across several learning modules, ECPs learn how to frame questions in ways that are open-ended and don't bias particular responses. Additionally, ECPs learn what questions they

should ask *themselves* when confronted with possible child abuse; and are shown the associated thought process that experts undertake when considering those questions (see Figure 4). ECPs also learn about the important distinction between (1) asking questions to learn more information, determine what was happening, and/or gain situational awareness, and (2) asking questions to 'investigate' whether abuse is occurring. *iLookOut* expressly teaches ECPs how good question-asking can help determine whether there is reasonable suspicion of abuse (which then triggers mandated reporting). Various learning modules also emphasize what information is needed to make an effective report; that 'certainty' is not required to make a report; and that as a mandated reporter a ECP's role is to report reasonable suspicions of child abuse, *not* to investigate whether abuse is in fact occurring.

Modeling

In both its video storyline and learning modules, *iLookOut* embraces Bandura's Social Model theory (Bandura, 1969), which posits that people often learn best through observation of others. For example, the interaction between Elisha and Megan (the lead characters in the *iLookOut* storyline) models how ECPs can rely on existing relationships for guidance when they are faced with challenging situations and don't know how to move forward. Learners also observe how questions can help them deal with doubts and uncertainty. Because such modelling is only helpful to the extent that learners can see how to apply it to their own situations, *iLookOut*'s storyline was expressly designed to reflect the real-life experiences of ECPs – in which key details about a child's situation are often unknown or unclear.

Safe environment

Research suggests that when dealing with complex subject matter, long-term learning is enhanced by providing practice opportunities for individuals to apply what they have learned, and do so in an

iLookOut: Keeping Track Wo... | Facts About Child Abuse: M... | Types of Child Abuse: Maine... | Red Flags for Child Abuse... | iLookOut: Keeping Track Wo...

v.lookoutmaine.org/home/student/1001/classroom/1001001/pathway/1002/trail/1/lesson/3/page/33

ib Slice Gallery | Imported From IE | Apple | Result List: the early... | Internet-based traini... | Latest Research | n... | center based care ch...

HOME | HELP | LOGOUT

iLookOut for Child Abuse

Resources

Pre-Training Check ☒ > Training ☒ > Post-Training Check ☐ > Learner Feedback ☐ > Next Steps ☐ > Completion ☐

A big part of being on the lookout involves knowing what questions need to be answered.

HOW OFTEN? Are you seeing a pattern of concerning injuries or behaviors?

HOW SEVERE? Is what you've observed "extreme" or within normal limits?

The **AGE** What is and isn't developmentally appropriate for a child this age?

The **EXPLANATION** Does the explanation "fit" what's happened?

The **IMPACT** How much is the child at risk and/or been harmed already?

Back Next

Figure 4. Interactive exercise to help learners ask better questions.

environment where learners feel safe to make mistakes (Knowles, Holton, & Swanson, 2012; Sheffler, 2009). *iLookOut* does this in part by not forcing learners into a binary, yes/no framework for considering child abuse. Rather, *iLookOut* gives learners opportunities to form judgments about various situations without the pressure of being right or wrong in each instance. This more iterative process offers learners a compelling, low-stakes learning environment in which they can try to predict what will (or should) happen, a practice space where they can 'try out' different perspectives and make mis-steps. In a safe environment, 'failure' prompts learners to reconsider their approach, and to problem-solve – which expands the scope of that learning environment. By creating an environment in which learners are encouraged (but not required) to explore additional content, take chances with their decision-making, and be exposed to realistic consequences for making a wrong or poor decision, *iLookOut* aims to better prepare ECPs for the uncertainty of real-world situations, where the decisions they make will affect the well-being of children.

Pinging

It is well established that one-time interventions typically fall victim to the Ebbinghaus forgetting curve – which shows that most gains in knowledge are lost within 10–21 days (Murre & Dros, 2015). To avoid this pitfall, learners who complete *iLookOut* can download the *iLookOut* app to receive weekly 5–7 min, gamified, micro-learning activities (referred to as 'pings') that reinforce learning and help ECPs stay current on best practices for protecting children. These pings take various forms (scenarios, animated videos, games, etc.), are interactive, and are sent to ECPs' smart phones, tablets, or laptops using a mobile-based technology. A back-end micro-learning platform then tracks learner responses, and provides professional development credit (at no cost) to ECPs who successfully complete these follow-up advanced training activities. A detailed description of *iLookOut* pinging is forthcoming. However, in brief, pings review and expand on key topics in the learning programme (e.g. red flags for abuse, when and how to act on one's concerns, etc.), as well as introduce new material (e.g. what is normal sexual development in young children) that can help ECPs decide how best to respond in a given situation. A separate research project is underway to evaluate the impact of pinging on ECPs' learning and behaviour.

Summary

The *iLookOut* learning programme offers an innovative and integrated approach for helping prepare early childhood professionals to identify and protect young children at risk for abuse. However, its design may be helpful as a prototype for a wide range of other educational interventions – particularly when the goals include cognitive, affective, and behavioural outcomes. Obviously, the resources needed for professional audio-visual and computer programming will typically be unavailable for small scale education projects. But for topics that have a wide audience of learners, the innovative, evidence-based pedagogical techniques described here may prove quite useful for enhancing early childhood professionals' engagement, learning, and motivation.

Note

1. For this paper, the term 'early childhood professionals' is used to denote early childhood educators, daycare providers, childcare workers, family and group home child care providers, pre-school teachers, and others who care for young children.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Eunice Kennedy Shriver National Institute of Child Health and Human Development [grant number 1 R01 HD088448-03].

Notes on contributor

Benjamin H. Levi is a Professor of Humanities and Pediatrics at the Penn State College of Medicine, and his research focuses on applied aspects of human decision-making.

Andrew Belser is a Professor of Theater at Penn State University, and Director of the Arts & Design Research Incubator.

Karl Kapp is a Professor of Instructional Technology at Bloomsburg University, and Director of the Institute for Interactive Technologies.

Nicole Verdiglione is the Project Manager of the iLookOut for Child Abuse learning program and research project.

Claudia Mincemoyer is a Professor Emerita of Sociology and Education at Penn State University, and former Director of Better Kid Care.

Sarah Dore is the Director of Learning and Curriculum for the iLookOut for Child Abuse learning program and research project.

Jane Keat is a Professor Emerita of Behavioral Sciences and Education at Penn State Harrisburg, and former Director of the Capital Area Early Childhood Training Institute.

Rick Fiene is a Professor Emeritus of Human Development and Family Systems and Psychology at Penn State University, and Director of the Research Institute for Key Indicators.

ORCID

Benjamin H. Levi  <http://orcid.org/0000-0003-2380-7109>

Jane Keat  <http://orcid.org/0000-0002-5021-7422>

References

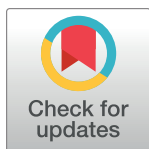
- Administration on Children, Youth and Families, US Department of Health and Human Services. (2016). *Child Maltreatment*. Washington, DC: US Government Printing Office.
- Bandura, A. (1969). Social-learning theory of identificatory processes. *Handbook of Socialization Theory and Research*, 213, 262.
- Boller, S., & Kapp, K. (2017). *Play to learn: Everything you need to know about designing effective learning games*. Alexandria, Virginia: Association for Talent Development.
- Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145–182.
- Dickinson, Emily. (1999). Tell all the truth but tell it slant. In R. W. Franklin (Ed.), *The poems of Emily Dickinson*. Cambridge, Massachusetts: Belknap Press of Harvard University Press.
- Dinehart, L. H., Katz, L. F., Manfra, L., & Ullery, M. A. (2013). Providing quality early care and education to young children who experience maltreatment: A review of the literature. *Early Childhood Education Journal*, 41(4), 283–290.
- Dixon, R. A., & Lammi, M. (2014). Cognitive mapping techniques: Implications for research in engineering and technology education. *Journal of Technology Education*, 25(2), 2–17.
- Ericsson, A., & Pool, R. (2016). *Peak: Secrets from the new science of expertise*. New York, New York: Houghton Mifflin Harcourt.
- Gibson, D., Ostashevski, N., Flintoff, K., Grant, S., & Knight, E. (2015). Digital badges in education. *Education and Information Technologies*, 20(2), 403–410.
- Kalyuga, S., & Sweller, J. (2018). Cognitive load and expert reversal. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 806–880). Cambridge, England, United Kingdom: Cambridge University Press.
- Kapp, K. M. (2012). *The gamification of learning and instruction*. San Francisco: Wiley.
- Kapp, K. M., Dore, S., Fiene, R., Verdiglione, N., Panlilio, C., Hamm, R. M., ... Levi, B. H. (2019). *Cognitive mapping for iLookOut for child abuse: An online training program for early childhood professionals*.
- Knowles, M. S., Holton, E. F., III, & Swanson, R. A. (2012). *The adult learner*. New York, New York, United States: Routledge.

- Levi, B. H., Dellasega, C., & Boehmer, S. (2012). What is reasonable suspicion of child abuse? *Journal of Public Child Welfare*, 6(5), 569–589.
- Levi, B., & Loeben, G. (2004). Index of suspicion: Feeling not believing. *Theoretical Medicine and Bioethics*, 25(4), 277–310.
- Levi, B. H., & Portwood, S. G. (2011). Reasonable suspicion of child abuse: Finding a common language. *The Journal of Law, Medicine & Ethics*, 39(1), 62–69.
- Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science*, 5(4), 333–369.
- Martin, J., & Powers, M. E. (1983). Truth or corporate propaganda: The value of a good war story. *Organizational Symbolism*, 1, 81–92.
- Mathews, B., Yang, C., Lehman, E. B., Mincemoyer, C., Verdiglione, N., & Levi, B. H. (2017). Educating early childhood care and education providers to improve knowledge and attitudes about reporting child maltreatment: A randomized controlled trial. *PLoS One*, 12(5), e0177777.
- McAdams, D. P. (2006). The role of narrative in personality psychology today. *Narrative Inquiry*, 16(1), 11–18.
- Murre, J. M., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLoS One*, 10(7), e0120644.
- National Association for Regulatory Administration. (2013). 2011–2013 child care licensing study. Retrieved from http://naralicensing.org/licensing_study
- Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: A systematic review and meta-analysis. *PLoS Medicine*, 9(11), e1001349.
- Panlilio, C., Famularo, L., Dore, S., Grable, B., Verdiglione, N., Yang, C., ... Levi, B. H. (2019). A mixed methods approach to investigating the structural and response process validity evidence for the iLookOut for child abuse knowledge test.
- Sheffler, J. L. (2009). Creating a warm and inclusive classroom environment: Planning for all children to feel welcome. *Electronic Journal for Inclusive Education*, 2(4), 4.
- Stadler, J., Dugmore, C., Venables, E., MacPhail, C., & Delany-Moretlwe, S. (2013). Cognitive mapping: Using local knowledge for planning health research. *BMC Medical Research Methodology*, 13(1), 96.
- Stott, A., & Neustaedter, C. (2013). Analysis of gamification in education. *Surrey, BC, Canada*, 8, 36.
- Whitebook, M., Phillips, D., & Howes, C. (2014). *Worthy work, STILL unlivable wages: The early childhood workforce 25 years after the national child care staffing study*. Berkeley, CA: Center for the Study of Child Care Employment.
- Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., ... Conckey, C. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming*, 40(2), 217–266.
- Yang, C., Panlilio, C., Verdiglione, N., Lehman, E. B., Hamm, R. M., Fiene, R., ... Levi, B. H. (2019). Generalized findings from a randomized controlled trial to a real-world study of an online education program on child abuse reporting.

RESEARCH ARTICLE

Generalizing findings from a randomized controlled trial to a real-world study of the *iLookOut*, an online education program to improve early childhood care and education providers' knowledge and attitudes about reporting child maltreatment

Chengwu Yang^{1*}, Carlo Panlilio², Nicole Verdiglione³, Erik B. Lehman⁴, Robert M. Hamm⁵, Richard Fiene^{6,7}, Sarah Dore³, David E. Bard⁸, Breanna Grable⁶, Benjamin Levi^{9*}



1 Department of Epidemiology and Health Promotion, New York University College of Dentistry, New York City, New York, United States of America, **2** Department of Educational Psychology, Counseling and Special Education, Pennsylvania State University, University Park, Hershey, Pennsylvania, United States of America, **3** Departments of Humanities, Penn State College of Medicine, Hershey, Pennsylvania, United States of America, **4** Departments of Population Health Sciences, Penn State College of Medicine, Hershey, Pennsylvania, United States of America, **5** Department of Family and Preventive Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, United States of America, **6** Departments of Psychology & Human Development Research Center, Pennsylvania State University, Middletown, Pennsylvania, United States of America, **7** Research Institute for Key Indicators, Middletown, Pennsylvania, United States of America, **8** Department of Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, United States of America, **9** Departments of Humanities & Pediatrics, Penn State College of Medicine, Hershey, Pennsylvania, United States of America

* chengwu.yang@nyu.edu (CY); bhlevi@psu.edu (BL)

OPEN ACCESS

Citation: Yang C, Panlilio C, Verdiglione N, Lehman EB, Hamm RM, Fiene R, et al. (2020) Generalizing findings from a randomized controlled trial to a real-world study of the *iLookOut*, an online education program to improve early childhood care and education providers' knowledge and attitudes about reporting child maltreatment. PLoS ONE 15 (1): e0227398. <https://doi.org/10.1371/journal.pone.0227398>

Editor: Andrew Carl Miller, East Carolina University Brody School of Medicine, UNITED STATES

Received: August 5, 2019

Accepted: December 13, 2019

Published: January 8, 2020

Copyright: © 2020 Yang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Funding for *iLookOut* was provided by Penn State University Center for the Protection of Children (BL), and NIH/NICHHD (National Institute of Child Health and Human Development) Grant Number 1 R01 HD088448-01 (BL, CY). The funders

Abstract

In recent years, real-world studies (RWS) are gaining increasing interests, because they can generate more realistic and generalizable results than randomized controlled clinical trials (RCT). In 2017, we published a RCT in 741 early childhood care and education providers (CCPs). It is the Phase I of our *iLookOut for Child Abuse* project (*iLookOut*), an online, interactive learning module about reporting suspected child maltreatment. That study demonstrated that in a RCT setting, the *iLookOut* is efficient at improving CCPs' knowledge of and attitudes towards child maltreatment reporting. However, the generalizability of that RCT's results in a RWS setting remains unknown. To address this question, we design and conduct this large RWS in 11,065 CCPs, which is the Phase II of the *iLookOut*. We hypothesize replication of the earlier RCT findings, i.e., the *iLookOut* can improve CCPs' knowledge of and attitudes toward child maltreatment reporting in a real world setting. In addition, this RWS also explores whether demographic factors affect CCPs' performance. Results of this RWS confirmed the generalizability of the previous RCT's results in a real world setting. It yielded similar effect sizes for knowledge and attitudes as were found in the earlier RCT. Cohen's d for knowledge improvement was 0.95 in that RCT, 0.96 in this RWS; Cohen's d for attitude improvement was 0.98 in that RCT, 0.80 in this RWS. Also, we found several

had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

significant differences in knowledge and attitude improvement with regard to age, race, education, and employment status. In conclusion, *iLookOut* improves knowledge and attitudes of CCPs about child maltreatment prevention and reporting in a real-world setting. The generalizability of the initial RCT findings to this RWS provides strong evidence that the *iLookout* will be effective in other real world settings. It can be a useful model for other interventions aimed at preventing child maltreatment.

Clinical trial registration for the original RCT: [NCT02225301](https://clinicaltrials.gov/ct2/show/study/NCT02225301) (ClinicalTrials.gov Identifier)

Introduction

While randomized controlled trials (RCT) have long been seen as the “gold standard” for evaluating the efficacy of interventions, there are well-known limitations to their generalizability [1]. Accordingly, there have been growing interests in real-world studies (RWS) to generate real-world evidence (RWE) that are more realistic and generalizable [2–9], and RWE is increasingly valued by regulators and payers [10]. In addition, RWE and the RCT can happily co-exist and complement each other [9].

Recently, we published data from an RCT about the online educational intervention, the *iLookOut for Child Abuse (iLookOut)*, showing that it improved early childhood care and education providers (CCPs) knowledge and attitudes about child maltreatment and its reporting [11]. In this follow-up study, through an RWS, we evaluate whether these results are generalizable to a broad population of CCPs in a real-world setting.

There are more than 675,000 confirmed cases of child maltreatment annually in the United States [12], but less than 1% of these are reported by CCPs (U.S. DHHS, 2017). This extremely low report rate by CCPs is alarming, given the fact that about 12 million U.S. children are served in some form of a child care setting, that children five years-old or younger account for 46% of confirmed maltreatment and more than 75% of maltreatment-related deaths (U.S. DHHS, 2017), and that the true incidence of child maltreatment is likely much higher than currently detected [13, 14]. Such underreporting suggests a need for CCPs to become better prepared to protect young children from maltreatment by improving their knowledge and attitude towards child maltreatment reporting. As has been identified by the Institute of Medicine and others, a key obstacle to improving awareness and reporting is the lack of evidence-based interventions [15–17]. In addition, the US Preventive Services Task Force (USPSTF) recently called for more evidence-based primary care interventions to prevent child maltreatment [12]. Several small studies have evaluated in-person training for CCPs [18, 19], and a brief online intervention [20, 21]. However, large studies involving scalable interventions are still lacking.

To meet this need, we created *iLookout*, an interactive online learning program designed specifically for CCPs (<https://ilookoutproject.org/>). An initial RCT using a test and re-test design with 741 participants demonstrated the feasibility of this three-hour online training, as well as its efficacy at increasing knowledge and changing attitudes about child maltreatment and its reporting [11]. Though this initial trial was promising, with large Cohen's *d* effect sizes for knowledge (0.95) and attitudes (0.98), its generalizability was limited by several factors, notably the potential for selection bias. Participants were enrolled only if the director of the child care program responded to the recruitment mailing. Family- and home-based CCP programs were under-represented, as were racial and ethnic minorities. In addition, enrollment was limited to a four-week period in early summer. Also, the sample size limited the opportunity for in-depth comparisons among subgroups.

To address these limitations, the present RWS used a statewide, open-enrollment design to enlist a larger, more representative sample of CCPs. We hypothesized that *iLookOut*'s efficacy at increasing knowledge and attitudes would be confirmed in this real-world sample, and our exploratory aim was to evaluate the impact of key demographic characteristics.

Materials and methods

Design

The Penn State College of Medicine Institutional Review Board approved this study prior to its initiation (IRB #: 1243). This RWS employed an open enrollment, single group, pre- and post-test design. Participants completed a demographic questionnaire, as well as previously validated knowledge and attitude measures regarding child maltreatment and its reporting [11]. Given the observational feature of this RWS, we have ensured that the manuscript adheres to the appropriate Equator Network guidelines, such as the STROBE (*Strengthening the Reporting of Observational Studies in Epidemiology*) Statement [22].

Participants

As an open-enrollment RWS, participants were not actively recruited to this study. However, all mandated reporters in Pennsylvania (including CCPs) are required by law to complete a mandated reporter training, and *iLookOut* was one of more than a dozen state-approved trainings listed on Pennsylvania's Department of Human Services website, and was available online at no charge. As such, online searches and word of mouth were the means for dissemination. Participant data reported here are from CCPs who completed *iLookOut* between January 2015 and March 2018. CCPs provided online informed consent prior to participating, and earned three hours of professional development credit for completing the learning program. No other incentives or remuneration were provided.

Intervention

The *iLookOut* online learning program uses an interactive, video-based storyline in which the learners take the role of a teacher of 4–5 year-olds at a child care facility. As key events unfold through interactions involving children, parents, and co-workers (all played by actors), the learners have to decide how to best respond. At different points, learners are posed questions. Based on their answer, they are provided didactic material to educate them about various aspects of child maltreatment. Other times, the learners must choose how to respond to events in the story. Throughout the learning program, CCPs can access multiple resource files covering definitions of maltreatment, facts about maltreatment, red flags, etc.[11].

Measures

The pre- and post-test comprise two parts. The first is a 21-item, true or false, expert-validated instrument previously described [11]. It measures individuals' knowledge about what constitutes child maltreatment, risk factors for maltreatment, and legal requirements for reporting suspected maltreatment. Correct answer to each of the 21 true or false items is scored as 1 point, and wrong answer is scored as 0 point. Therefore, the total score of the knowledge scale ranges from 0 to 21, which higher score representing more knowledge about child maltreatment. The second part contains 13 items, rated on 7-point Likert-style scales, from a previously validated instrument [23] adapted to comport with Pennsylvania jurisdictional standards. It measures individuals' attitude towards reporting potential child maltreatment. An individual's attitude score is the average score of the 13 items, ranges from 1 to 7, with higher score

representing more positive attitude towards reporting potential child maltreatment. The pre- and post-test question items were identical, but to minimize recall bias, their sequencing orders were changed between the pre- and the post-test.

Sample size and statistical analysis

Given the RWS nature of this study, no a priori sample size estimation was planned. However, post-hoc power analyses were implemented to check the statistical power for some important subgroup analyses [4]. We also compared participant demographics between the initial RCT and this RWS.

As with the RCT, the statistical analysis of this RWS examined *iLookOut*'s impact on CCPs' knowledge and attitudes related to child maltreatment and its reporting. The two primary outcome variables were the total knowledge score and the total attitude score, both measured as "change", i.e., total score at post-test minus at pre-test. The analysis focused on whether the present RWS confirmed the results of the initial RCT. To compare effect sizes between the RCT and the RWS, we used two measures: 1) the absolute difference, i.e., the measured change in pre- to post-test score for the RWS, minus the measured change in initial RCT; and 2) the Cohen's d calculation [24]. In addition, we explored the impact of demographic factors on these two primary outcome variables through analysis of covariance (ANCOVA), framing demographic variables as covariates, and adjusting for pre-measurement scores. These demographic variables include age, gender, race/ethnicity, education, employment, parent/guardian status, prior trained status, work environment, years as practitioner, primary job responsibilities, and religiosity. We used the SAS software package, version 9.4, for statistical analyses, and the G*Power software package, version 3.1.9, for post-hoc power analyses.

Results

During the 38 months of the RWS reported here, 11,605 CCPs completed the *iLookOut* online training. Compared to those CCPs in the initial RCT, these RWS participants were more representative of the general population of CCPs in Pennsylvania, particularly for its enrollment of Blacks (20.8% vs. 8.0%) and males (10.9% vs. 2.3%). In addition, the CCPs in this RWS were younger (48.0% vs. 40.4% aged below 30), and a greater proportion worked in more urban area (36.4% vs. 22.1%). Table 1 illustrates comparisons of full demographics between these two studies.

Table 2 illustrates comparisons of the *iLookOut* training's effect sizes on knowledge and attitude scores between this RWS and the RCT, demonstrating improved knowledge and attitudes about child maltreatment reporting for both studies. Pre- to post- changes in knowledge score increased by 2.80 for RWS participants, compared to 2.65 in the initial RCT, a 5.7% relative change. The Cohen's d on the total knowledge score was 0.96 in this RWS versus 0.95 in the RCT, a 1% relative change. The pre-to post- change in attitude average score was 0.5 for RWS participants, versus 0.59 in the initial RCT, a -15.3% relative change. The Cohen's d on the average attitude score was 0.80 in this RWS, versus 0.98 in the RCT, a relative change of -18.4%.

Table 3 summarizes the results of exploratory multivariate analyses (ANCOVA) for each of the two outcome variables (knowledge and attitude scores) with all of the demographic variables. After adjustment for pre-measurement scores and all the other demographic variables, only four demographics (age, race, education, and employment) showed impacts on either of the two outcome variables, with age and education being positively correlated with increase in knowledge scores.

Table 1. Comparisons of demographic characteristics of early childcare professionals.

		Phase II: RWS	Phase I: RCT	Difference	p-value
Sample Size		11,065	741	10,324	
Age	18–29	5309 (48.0%)	299 (40.4%)	7.6%	<0.001
	30–44	2912 (26.3%)	216 (29.1%)	-2.8%	
	45+	2844 (25.7%)	226 (30.5%)	-4.8%	
Gender	Male	1210 (10.9%)	17 (2.3%)	8.6%	<0.001
	Female	9855 (89.1%)	724 (97.7%)	-8.6%	
Race/Ethnicity	Non-Hispanic White	7605 (68.7%)	624 (84.2%)	-15.5%	<0.001
	Non-Hispanic Black	2296 (20.8%)	59 (8.0%)	12.8%	
	Hispanic	658 (6.0%)	25 (3.4%)	2.6%	
	Asian	227 (2.1%)	15 (2.0%)	0.1%	
	Other	279 (2.4%)	18 (2.4%)	0.0%	
Education	Below High School	82 (0.7%)	0 (0.0%)	0.7%	<0.001
	High School or GED	4611 (41.7%)	197 (26.6%)	15.1%	
	Child Development Associate (CDA)	765 (6.9%)	101 (13.6%)	-6.7%	
	Associates	1483 (13.4%)	149 (20.1%)	-6.7%	
	Bachelors	2983 (27.0%)	229 (30.9%)	-3.9%	
	Masters or Doctoral	1141 (10.3%)	65 (8.8%)	1.5%	
Employment	Permanent Full-Time	6276 (56.7%)	534 (72.1%)	-15.4%	<0.001
	Permanent Part-Time	2943 (26.6%)	169 (22.8%)	3.8%	
	Contract for special services	177 (100.0%)	0 (0.0%)	100.0%	
	Substitute Teacher	206 (1.9%)	6 (0.8%)	1.1%	
	Seasonal	793 (7.2%)	28 (3.8%)	3.4%	
	Volunteer	334 (3.0%)	0 (0.0%)	3.0%	
	Other	336 (4.6%)	4 (0.5%)	4.1%	
Parent/Guardian	Yes	6089 (55.0%)	452 (61.0%)	-6.0%	0.002
	No	4976 (45.0%)	289 (39.0%)	6.0%	
Prior Trained	Yes	7371 (66.6%)	582 (78.5%)	-11.9%	<0.001
	No	3694 (33.4%)	159 (21.5%)	11.9%	
Work Environment	Rural	2191 (19.8%)	206 (27.8%)	-8.0%	
	Suburban	4848 (43.8%)	371 (50.1%)	-6.3%	<0.001
	Urban	4026 (36.4%)	164 (22.1%)	14.3%	
Years as Practitioner	<1	3272 (29.9%)	68 (9.2%)	20.7%	
	1–2	1652 (14.9%)	112 (15.1%)	-0.2%	<0.001
	3–5	2034 (18.4%)	145 (19.6%)	-1.2%	
	6–10	1657 (15.0%)	154 (20.8%)	-5.8%	
	11–15	887 (8.0%)	75 (10.1%)	-2.1%	
	>15	1563 (14.1%)	187 (25.2%)	-11.1%	
Primary job responsibilities	Teacher/caregiving staff (infant–grade 4)	7049 (63.7%)	555 (75.0%)	-11.3%	
	Early intervention specialist	184 (1.7%)	0 (0.0%)	1.7%	<0.001
	Support staff	651 (5.9%)	25 (3.4%)	2.5%	
	Director/Assistant Director	781 (7.1%)	95 (12.8%)	-5.7%	
	Other	2400 (21.7%)	66 (8.8%)	12.9%	
Religiosity	Extremely unreligious	198 (1.8%)	10 (1.4%)	0.4%	
	Unreligious	689 (6.2%)	54 (7.4%)	-1.2%	<0.001
	Somewhat unreligious	436 (3.9%)	13 (1.8%)	2.1%	
	Neutral	2568 (23.2%)	117 (15.9%)	7.3%	
	Somewhat religious	2503 (22.6%)	215 (28.8%)	-6.2%	

(Continued)

Table 1. (Continued)

		Phase II: RWS	Phase I: RCT	Difference	p-value
	Religious	4017 (36.3%)	287 (38.9%)	-2.6%	
	Extremely religious	654 (5.9%)	45 (5.9%)	0.0%	

<https://doi.org/10.1371/journal.pone.0227398.t001>

Post hoc power analysis indicates that with a total sample size of 11,065, at an alpha level of 0.05, with 80% power, the ANCOVA with 11 covariates would be able to detect an effect size as small as 0.04 among six groups; and using an effect size cut-off of 0.25, then the power would approach to 99.5%.

Discussion

The results from this RWS demonstrate that in a large, representative sample of child care professionals (CCPs), the online *iLookOut* learning program is effective at improving knowledge and changing attitudes about child maltreatment and its reporting. These findings confirm the conclusions from the initial RCT of *iLookOut*, and demonstrate the feasibility of scaling this evidence-based, online mandated reporter training. This is notable insofar as more than 11,000 CCPs completed *iLookOut*, even when no special incentives were offered, and they reported being highly satisfied with the learning experience (paper forthcoming). No significant differences were identified with regard to CCPs' parenting status, previous training, work environment, years as practitioner, primary job responsibility, or religiosity. However, age, race, education, and employment affected changes in knowledge or attitude scores, with older and more educated CCPs achieving increased gains in knowledge scores.

The generalizability of the initial RCT findings provides supporting evidence that the *iLookOut* online learning program will be effective in other real world settings, and may be a useful model for other interventions aimed at preventing child maltreatment [12]. *iLookOut*'s general storyline and overall format are generalizable for all kinds of CCPs in all U.S. states, in part because state-specific information is housed in discrete learning modules (within the learning program) that can be readily adapted to comport with the laws and policies of different states. The efficacy of *iLookOut* does not appear to be affected by previous training, work environment, years as practitioner, primary job responsibility, parenting status, or religiosity. However, larger gains in knowledge were seen in CCPs who were older, more highly educated, employed seasonally, or white. More research is warranted to better understand the underpinnings of these differences, and how best to optimize gains in knowledge for all CCPs.

The statistical analyses reported here focus on effect sizes, instead of p-values, for several reasons. First, p-values are not a good measure of evidence [25]. Second, the misuse and maltreatment of p-values has led both researchers and the American Statistical Association to raise concerns about the limitations of p-value-driven conclusions [26, 27]. Third, the very large sample size (over 11,000) of this RWS could yield findings of statistical significance for even very small effect sizes that have no clinical significance [28]. Fourth, the large difference in

Table 2. Comparisons of effect sizes on knowledge and attitude scores.

	Post-Pre in Phase II: RWS	Post-Pre in Phase I: RCT	Difference: Phase II-Phase I	Relative Change	Cohen's d in Phase II: RWS	Cohen's d in Phase I: RCT	Difference: Phase II-Phase I	Relative Change
Knowledge: Total Score (Range: 0–21)	2.80 ± 2.90	2.65 ± 2.78	0.15	5.7%	0.96	0.95	0.01	1.0%
Attitude: Average Score (Range: 1–7)	0.50 ± 0.63	0.59 ± 0.60	-0.09	-15.3%	0.80	0.98	-0.18	-18.4%

<https://doi.org/10.1371/journal.pone.0227398.t002>

Table 3. Summary of analysis of covariance (ANCOVA) results.

Variable**	Total Knowledge Score (0–21)			Average Attitude Score (1–7)		
	Pre (Mean ± SD)	Post (Mean ± SD)	Mean Change (95% CI)*	Pre (Mean ± SD)	Post (Mean ± SD)	Mean Change (95% CI)*
Age						
18–29	13.7 ± 2.7	16.6 ± 2.9	2.9 (2.8, 2.9)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
30–44	14.1 ± 3.0	17.2 ± 2.9	3.2 (3.1, 3.3)	5.9 ± 0.8	6.4 ± 0.7	0.5 (0.5, 0.5)
More than 44	13.9 ± 3.1	17.4 ± 2.8	3.5 (3.4, 3.6)	5.8 ± 0.8	6.4 ± 0.7	0.5 (0.5, 0.5)
Gender						
Male	13.8 ± 3.0	16.9 ± 3.0	3.0 (2.8, 3.1)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.4, 0.5)
Female	13.9 ± 2.9	16.9 ± 2.9	3.1 (3.1, 3.2)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Race						
White	14.0 ± 2.9	17.4 ± 2.8	3.5 (3.4, 3.5)	5.9 ± 0.7	6.4 ± 0.6	0.6 (0.5, 0.6)
Black or African American	13.5 ± 2.9	15.8 ± 2.9	2.2 (2.1, 2.3)	5.8 ± 0.8	6.2 ± 0.8	0.3 (0.3, 0.4)
American Indian or Alaska Native	13.5 ± 3.2	16.1 ± 2.6	2.7 (1.9, 3.5)	6.1 ± 0.8	6.3 ± 0.7	0.4 (0.2, 0.5)
Hispanic	13.3 ± 2.8	16.0 ± 2.8	2.5 (2.4, 2.7)	5.9 ± 0.8	6.3 ± 0.7	0.4 (0.4, 0.5)
Asian	13.2 ± 3.4	16.0 ± 3.3	2.0 (1.7, 2.4)	5.8 ± 0.8	6.3 ± 0.7	0.4 (0.4, 0.5)
Native Hawaiian/ Pacific Islander	13.6 ± 3.6	16.1 ± 3.1	2.3 (1.1, 3.5)	5.6 ± 1.0	6.1 ± 1.1	0.4 (0.2, 0.7)
Other	13.2 ± 3.4	16.4 ± 3.3	2.9 (2.5, 3.2)	5.8 ± 0.8	6.3 ± 0.8	0.5 (0.4, 0.6)
Education						
8th Grade	12.8 ± 3.4	15.2 ± 3.2	1.7 (1.1, 2.2)	5.6 ± 0.9	6.0 ± 0.9	0.2 (0.1, 0.4)
High School Diploma or G.E.D.	13.4 ± 2.9	16.3 ± 3.0	2.6 (2.6, 2.7)	5.8 ± 0.8	6.3 ± 0.7	0.4 (0.4, 0.5)
Child Development Associate	14.0 ± 3.1	16.4 ± 2.9	2.6 (2.4, 2.8)	5.8 ± 0.8	6.3 ± 0.7	0.5 (0.4, 0.5)
Associate's Degree	13.9 ± 2.9	16.8 ± 2.7	2.9 (2.8, 3.1)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Bachelor's Degree	14.2 ± 2.9	17.8 ± 2.7	3.8 (3.7, 3.9)	5.9 ± 0.7	6.5 ± 0.6	0.6 (0.6, 0.6)
Masters or Doctoral Degree	14.5 ± 3.0	18.0 ± 2.7	3.9 (3.7, 4.0)	6.0 ± 0.7	6.5 ± 0.6	0.6 (0.6, 0.6)
Employment						
Permanent full-time	14.0 ± 2.9	16.9 ± 2.9	3.0 (2.9, 3.0)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Permanent part-time	13.5 ± 2.9	16.7 ± 3.0	3.1 (3.0, 3.2)	5.9 ± 0.8	6.3 ± 0.7	0.5 (0.5, 0.5)
Contract for special services/care	14.8 ± 2.4	18.0 ± 2.5	3.3 (2.9, 3.7)	6.0 ± 0.7	6.5 ± 0.6	0.5 (0.4, 0.6)
Substitute teacher	13.6 ± 3.1	17.3 ± 2.9	3.5 (3.1, 3.9)	6.0 ± 0.7	6.5 ± 0.7	0.5 (0.5, 0.6)
Seasonal or short-term	13.6 ± 2.7	17.4 ± 2.8	3.9 (3.7, 4.1)	5.9 ± 0.7	6.5 ± 0.6	0.6 (0.6, 0.7)
Volunteer	13.7 ± 3.2	17.7 ± 2.8	3.7 (3.4, 4.0)	5.9 ± 0.7	6.5 ± 0.6	0.6 (0.5, 0.6)
Other	13.5 ± 3.1	17.2 ± 2.9	3.4 (3.2, 3.7)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.4, 0.6)
Parent or guardian of child						
Yes	14.0 ± 3.0	17.0 ± 2.9	3.0 (3.0, 3.1)	5.9 ± 0.8	6.4 ± 0.7	0.5 (0.5, 0.5)
No	13.7 ± 2.8	16.9 ± 3.0	3.2 (3.1, 3.3)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Previously trained						
Yes	14.2 ± 2.8	17.1 ± 2.8	3.1 (3.0, 3.1)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
No	13.1 ± 2.9	16.6 ± 3.0	3.2 (3.1, 3.3)	5.8 ± 0.8	6.3 ± 0.7	0.5 (0.5, 0.5)
Work Environment						
Rural	14.1 ± 2.9	17.1 ± 2.8	3.2 (3.1, 3.3)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Suburban	13.8 ± 3.0	17.2 ± 2.9	3.3 (3.2, 3.4)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Urban	13.7 ± 2.9	16.5 ± 3.0	2.9 (2.8, 3.0)	5.9 ± 0.8	6.3 ± 0.7	0.5 (0.5, 0.5)
Years as practitioner						
Less than 1	13.6 ± 2.8	16.9 ± 2.9	3.2 (3.1, 3.3)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
1–2	13.6 ± 2.9	16.7 ± 3.0	3.1 (3.0, 3.3)	5.9 ± 0.7	6.3 ± 0.7	0.5 (0.5, 0.5)
3–5	13.9 ± 2.9	16.8 ± 2.9	3.0 (2.9, 3.1)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)

(Continued)

Table 3. (Continued)

Variable**	Total Knowledge Score (0–21)			Average Attitude Score (1–7)		
	Pre (Mean ± SD)	Post (Mean ± SD)	Mean Change (95% CI)*	Pre (Mean ± SD)	Post (Mean ± SD)	Mean Change (95% CI)*
6–10	14.0 ± 3.0	17.1 ± 2.9	3.2 (3.0, 3.3)	5.8 ± 0.8	6.4 ± 0.7	0.5 (0.5, 0.5)
11–15	13.9 ± 3.2	17.0 ± 3.0	3.0 (2.8, 3.2)	5.8 ± 0.8	6.4 ± 0.7	0.5 (0.5, 0.5)
More than 15	14.3 ± 3.0	17.5 ± 2.8	3.0 (2.9, 3.2)	5.9 ± 0.8	6.4 ± 0.7	0.5 (0.5, 0.5)
Primary Job Responsibility						
Teacher/caregiving staff (age 0–5)	13.7 ± 2.9	16.8 ± 2.9	3.1 (3.1, 3.2)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Early intervention specialist	14.7 ± 2.8	17.6 ± 2.9	2.9 (2.5, 3.3)	6.0 ± 0.7	6.4 ± 0.7	0.4 (0.3, 0.5)
Kindergarten teacher	13.1 ± 3.0	16.6 ± 3.6	2.9 (2.5, 3.4)	5.7 ± 0.7	6.2 ± 0.8	0.4 (0.3, 0.5)
Early elementary teacher	13.5 ± 2.8	16.8 ± 3.1	3.1 (2.8, 3.3)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.6)
Support staff	13.6 ± 2.8	16.6 ± 2.9	3.0 (2.8, 3.2)	5.8 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.6)
Assistant Director	14.5 ± 2.9	17.2 ± 2.9	3.0 (2.7, 3.3)	5.9 ± 0.8	6.4 ± 0.6	0.5 (0.4, 0.6)
Director	15.0 ± 2.8	17.8 ± 2.6	3.1 (2.7, 3.3)	6.0 ± 0.7	6.5 ± 0.6	0.5 (0.4, 0.5)
Other	14.0 ± 2.9	17.3 ± 2.9	3.1 (3.0, 3.2)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Religiosity						
Extremely Unreligious	14.4 ± 2.9	17.4 ± 3.0	3.3 (2.9, 3.6)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.4, 0.5)
Unreligious	13.9 ± 3.1	17.0 ± 2.9	3.1 (2.9, 3.3)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.6)
Somewhat unreligious	14.1 ± 2.9	16.9 ± 2.8	3.0 (2.8, 3.3)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Neutral	13.7 ± 2.9	16.5 ± 3.0	2.9 (2.8, 3.0)	5.9 ± 0.8	6.3 ± 0.7	0.5 (0.5, 0.5)
Somewhat religious	13.9 ± 2.9	17.0 ± 2.9	3.1 (3.0, 3.2)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.5, 0.5)
Religious	13.8 ± 3.0	17.0 ± 2.9	3.2 (3.1, 3.2)	5.9 ± 0.7	6.4 ± 0.7	0.5 (0.4, 0.5)
Extremely religious	13.9 ± 2.9	17.8 ± 2.8	3.7 (3.5, 3.9)	5.9 ± 0.7	6.5 ± 0.7	0.5 (0.5, 0.6)

* The mean changes come from a multivariable model for the change in the outcome adjusted for the pre-measurement and including all of the following demographic variables as covariates: age, gender, race/ethnicity, education, employment, parent/guardian status, prior trained status, work environment, years as practitioner, primary job responsibilities, and religiosity. As a result, the mean changes displayed are adjusted for all of the other variables.

** All of the variables have p-values less than 0.05, except for gender (p = 0.061), and primary job responsibilities (p = 0.641).

<https://doi.org/10.1371/journal.pone.0227398.t003>

sample size between the initial RCT and this RWS renders effect sizes a more meaningful comparison than p-values. Finally, for proposed sub-group analyses involving many demographic covariates, p-values are less likely to yield meaningful findings [29]. Accordingly, we compared effect sizes by examining the overlap of their confidence limits.

The present findings are limited by potential biases encountered in all RWS, including selection bias, information bias, and confounding [3]. Multivariate analysis (ANCOVA) was used to try to account for these factors, and the initial RCT does provide additional reassurance that the present findings are valid. However, without qualitative data, an explanatory model for the present findings will remain incomplete.

Conclusion

This real-world study of more than 11,000 early childhood professionals (CCPs) who were neither recruited nor incentivized to complete the *iLookOut for Child Maltreatment* confirms that *iLookOut* significantly improves knowledge and attitudes regarding child maltreatment and its reporting. These results provide strong evidence that interactive, online interventions for helping prevent child maltreatment are both effective and scalable. A 5-year randomized controlled trial (<https://clinicaltrials.gov/ct2/show/NCT02225301?term=NCT02225301&rank=1>) is

currently underway to evaluate how well *iLookOut* helps CCPs identify and report true child maltreatment.

Supporting information

S1 Dataset.

(CSV)

Acknowledgments

We are grateful to all study participants and their institutions; to the Pennsylvania early childhood education and care agencies which facilitated the research; to participants and expert reviewers in the pilot phases; to the Center for the Application of Information Technologies (CAIT) for technological support; and to the Penn State University Center for the Protection of Children for general support.

Author Contributions

Conceptualization: Chengwu Yang, Benjamin Levi.

Data curation: Nicole Verdiglione, Erik B. Lehman, Sarah Dore, Breanna Grable, Benjamin Levi.

Formal analysis: Chengwu Yang, Erik B. Lehman, Benjamin Levi.

Funding acquisition: Benjamin Levi.

Investigation: Chengwu Yang, Carlo Panlilio, Nicole Verdiglione, Erik B. Lehman, Robert M. Hamm, Richard Fiene, Sarah Dore, David E. Bard, Benjamin Levi.

Methodology: Chengwu Yang, Robert M. Hamm, Benjamin Levi.

Project administration: Nicole Verdiglione, Sarah Dore, Breanna Grable, Benjamin Levi.

Resources: Benjamin Levi.

Supervision: Chengwu Yang, Benjamin Levi.

Validation: Chengwu Yang, Erik B. Lehman.

Visualization: Chengwu Yang, Erik B. Lehman.

Writing – original draft: Chengwu Yang, Carlo Panlilio, Richard Fiene, Benjamin Levi.

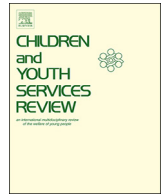
Writing – review & editing: Chengwu Yang, Carlo Panlilio, Nicole Verdiglione, Erik B. Lehman, Robert M. Hamm, Richard Fiene, Sarah Dore, David E. Bard, Breanna Grable, Benjamin Levi.

References

1. Kaptchuk TJ. The double-blind, randomized, placebo-controlled trial: gold standard or golden calf? *J Clin Epidemiol*. 2001; 54(6):541–9. [https://doi.org/10.1016/s0895-4356\(00\)00347-4](https://doi.org/10.1016/s0895-4356(00)00347-4) PMID: 11377113
2. Garrison LP Jr, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: The ISPOR real-world data task force report. *Value in Health*. 2007; 10(5):326–35. <https://doi.org/10.1111/j.1524-4733.2007.00186.x> PMID: 17888097
3. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-World Evidence—What Is It and What Can It Tell Us? *N Engl J Med*. 2016; 375(23):2293–7. <https://doi.org/10.1056/NEJMs1609216> PMID: 27959688

4. Administration USFD. Use of real-world evidence to support regulatory decision-making for medical devices. Guidance for industry and food and drug administration staff. Docket Number: FDA-2016-D-2153. 2017.
5. Sun X, Tan J, Tang L, Guo JJ, Li X. Real world evidence: experience and lessons from China. *BMJ*. 2018; 360:j5262. <https://doi.org/10.1136/bmj.j5262> PMID: 29437644
6. Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. *JAMA*. 2018; 320(9):867–8. <https://doi.org/10.1001/jama.2018.10136> PMID: 30105359
7. Food U, Administration D. Framework for FDA's real-world evidence program. 2018.
8. Klonoff DC. The Expanding Role of Real-World Evidence Trials in Health Care Decision Making. *Journal of Diabetes Science and Technology*. 2019;1932296819832653.
9. Evans K. Real World Evidence: Can We Really Expect It to Have Much Influence? *Drugs—Real World Outcomes*. 2019; 6(2):43–5. <https://doi.org/10.1007/s40801-019-0155-3> PMID: 31016548
10. Calvert MJ, O'Connor DJ, Basch EM. Harnessing the patient voice in real-world evidence: the essential role of patient-reported outcomes. *Nature Reviews Drug Discovery*. 2019; 18(5).
11. Mathews B, Yang C, Lehman EB, Mincemoyer C, Verdiglione N, Levi BH. Educating early childhood care and education providers to improve knowledge and attitudes about reporting child maltreatment: A randomized controlled trial. *PLoS ONE*. 2017; 12(5):e0177777. <https://doi.org/10.1371/journal.pone.0177777> PMID: 28542285
12. Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, Davidson KW, et al. Interventions to prevent child maltreatment: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2018; 320(20):2122–8. <https://doi.org/10.1001/jama.2018.17772> PMID: 30480735
13. Kohl PL, Jonson-Reid M, Drake B. Time to leave substantiation behind: Findings from a national probability study. *Child Maltreatment*. 2009; 14(1):17–26. <https://doi.org/10.1177/1077559508326030> PMID: 18971346
14. Stoltenborgh M, Bakermans-Kranenburg MJ, van IJzendoorn MH, Alink LR. Cultural–geographical differences in the occurrence of child physical abuse? A meta-analysis of global prevalence. *International Journal of Psychology*. 2013; 48(2):81–94. <https://doi.org/10.1080/00207594.2012.697165> PMID: 23597008
15. Alvarez KM, Kenny MC, Donohue B, Carpin KM. Why are professionals failing to initiate mandated reports of child maltreatment, and are there any empirically based training programs to assist professionals in the reporting process? *Aggression and Violent Behavior*. 2004; 9(5):563–78.
16. Dinehart L, Kenny MC. Knowledge of child abuse and reporting practices among early care and education providers. *Journal of Research in Childhood Education*. 2015; 29(4):429–43.
17. IOM, NRC. New directions in child abuse and neglect research. The National Academies Press Washington, DC; 2014.
18. McGrath P, Cappelli M, Wiseman D, Khalil N, Allan B. Teacher awareness program on child abuse: a randomized controlled trial. *Child Abuse & Neglect*. 1987; 11(1):125–32.
19. Khan A, Rubin D, Winnik G. Evaluation of the mandatory child abuse course for physicians: do we need to repeat it? *Public Health*. 2005; 119(7):626–31. <https://doi.org/10.1016/j.puhe.2004.10.010> PMID: 15925678
20. Kenny MC. Web-based training in child maltreatment for future mandated reporters. *Child Abuse & Neglect*. 2007; 31(6):671–8.
21. Kenny MC, Lopez-Griman AM, Donohue B. Development and initial evaluation of a cost-effective, internet-based program to assist professionals in reporting suspected child maltreatment. *Journal of Child & Adolescent Trauma*. 2017; 10(4):385–93.
22. Von Elm E, Altman D, Egger M, Pocock S, G? tzsche PC, Vandenbroucke JP: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med*. 2007; 4(10):e296. <https://doi.org/10.1371/journal.pmed.0040296> PMID: 17941714
23. Walsh K, Rassafiani M, Mathews B, Farrell A, Butler D. Exploratory factor analysis and psychometric evaluation of the teacher reporting attitude scale for child sexual abuse. *J Child Sex Abus*. 2012; 21(5):489–506. <https://doi.org/10.1080/10538712.2012.689423> PMID: 22994689
24. Cohen J. A power primer. *Psychological bulletin*. 1992; 112(1):155. <https://doi.org/10.1037/0033-2909.112.1.155> PMID: 19565683
25. Piantadosi S. Clinical trials: a methodologic perspective: John Wiley & Sons; 2017.
26. Nuzzo R. Scientific method: statistical errors. *Nature News*. 2014; 506(7487):150.

27. Wasserstein RL, Lazar NA. The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician*. 2016; 70(2):129–33.
28. Yang C, Vrana KE. Rescuing Suboptimal Patient-Reported Outcome Instrument Data in Clinical Trials: A New Strategy. *Healthcare (Basel)*. 2018; 6(1).
29. Rothwell PM. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *The Lancet*. 2005; 365(9454):176–86.



Identifying predictive indicators: The state of Washington foster care home study

Sonya Stevens^{a,*}, Richard Fiene^b, Daniel Blevins^a, Amber Salzer^a

^a Washington State Department of Children, Youth, and Families, Olympia, WA, United States

^b The Pennsylvania State University, Research Institute for Key Indicators and National Association for Regulatory Administration, United States

ARTICLE INFO

Keywords:

Foster care
Licensing
Home study
Differential
Predictive indicators
Kinship

ABSTRACT

A mixed method correlational exploratory pilot was conducted in Washington State to determine items within the home study assessment that could be used as indicators to identify baseline requirements of the assessment and suggest anticipated depth (expansion or reduction) within the required topic(s). The purpose of the home study is to assess the caregiver(s)' ability to provide a safe home, the quality of care needed by children and an environment that is nurturing, respectful and supportive. The goal of this study is to identify predictive indicators that will assist in the development of a home study that will increase consistency within home studies and decrease timeliness of completion.

The use of predictive indicators may have the potential to reduce subjective decision making as well as identify inconsistencies when determining the recommendation of approval or denial of a home study. Additionally, with a carefully designed home study system inclusive of predictive analytics, it is possible to reduce the amount of time an assessor uses to approve or deny a home study, saving agency time and resources. Finally, by using focused technical assistance with those applicants who need more or specific support, the use of predictive indicators may increase the success of timely placement and permanency goals. This mixed method study included a case review of 207 home studies where 19 primary and secondary themes emerged as significant. It lays the ground work for methods used to identify predictive elements within the assessment process. Preliminary results are provided along with further recommendations.

1. Introduction

The field of child welfare unites around three major goals for children, youth, and families: safety, permanency, and well-being (Conradi, Landsverk, & Wotring, 2014). Nationwide, when a family wants to become a foster home, they are required to meet individual state requirements ensuring a child is safe and well cared for. Requirements around foster care licensing are designed to reduce predictable risks to the health, safety, and well-being of children placed in foster homes (Cuccaro-Alamin, Foust, Vaithianathan, & Putnam-Hornstein, 2017). These requirements, or state laws and policies vary widely from state to state (Gateway, 2018). Yet, despite the many years of home study practice and state and federal requirements mandating the use of home study, the home study tool itself has received little attention in the research world. With such limited research surrounding the home study process it is yet unknown how the home study process can be completed using tangible data collection and analyses in addition to the professional clinical judgment commonly used (Crea, Barth, & Chintapalli,

2007).

The variation in regulations combined with the vast differences in home study assessors' background and training often leads to assessment results that vary greatly (DePanfilis & Girvin, 2005; Rossi, Schuerman, & Budde, 1999). Rossi et al. (1999) conducted a study using regression analysis and found that while assessors utilized the same characteristics when making decisions, the decisions themselves varied greatly. More recently, some agencies have begun to employ various risk assessment tools throughout child welfare to improve decision making of child removal and placement into out of home care (Cuccaro-Alamin et al., 2017). However, Cuccaro-Alamin et al. (2017), highlight the fact that while standardized tools are often more effective than simple clinical judgement, there are also multiple operational and statistical limitations to using those tools including the tool's validity and reliability, the usability and cost, limited accuracy, and inconsistent use amongst others.

Washington state, like other states and countries employs the use of clinical tools when considering foster care licenses and placement. In

* Corresponding author.

E-mail address: Sonya.stevens@dcyf.wa.gov (S. Stevens).

<https://doi.org/10.1016/j.childyouth.2020.105133>

Received 25 March 2020; Received in revised form 2 June 2020; Accepted 2 June 2020

Available online 06 June 2020

0190-7409/ Published by Elsevier Ltd.

Washington state, once a child is identified as needing out-of-home placement, removal is determined by law enforcement and/or court order, as recommended by Child Protective Services. Current policy states all placements must complete a home study (Department of Children, Youth, and Families, 2008). Home studies are completed through the Department of Children, Youth and Families (DCYF) Licensing Division. The purpose of the home study is to assess the caregiver(s)' ability to provide a safe home, the quality of care needed by children and an environment that is nurturing, respectful and supportive (Washington Administrative Code 110-148-1320, 110-148-1365).

If a child is placed in general foster care, the provider will have already completed a home study. The home study should be completed within 120 days of assignment. However, a child may be placed in kinship care (defined as a relative or suitable other) or in general foster care. Kinship care providers may elect to become a licensed foster parent, but if they choose not to, they are considered "unlicensed." When a child is placed in kinship care, a home study referral should be made to Licensing Division within 30 days of out-of-home placement. The goal of this study is to identify elements within a home study that are indicators that will assist in the development of a home study that could increase consistency and decrease timeliness of completion.

At any given time, there are approximately 1500 pending applications assigned to Licensing Division staff. Each application requires a completed home study unless the applicant chooses to withdraw from the process. In addition, there are approximately 400 un-referred home studies that will increase the pending workload once referred. The un-referred home studies result from identification of children who have been placed in out-of-home care, but a referral has not yet been made to Licensing Division. Due to the high volume of pending applications and the amount of time it takes to complete the home study, there has been a long standing backlog (pending over 120 days) of incomplete home studies. As of December 2019, there were a total 627 (42% of the total pending application) pending for over 120 days.

Due to the increasing issues surrounding the backlog of home studies, Licensing Division assessed the home study forms and processes for efficiency and consistency. During the review, it was important to consider effectiveness in conjunction with efficiency within the family assessment tools. The problem is there is not a clear understanding of what effectiveness means, in terms of the home study. Broad goals of safety, permanency and well-being of children in foster care are often quoted without a clear indication of what specific items lead to adverse determinations. The purpose of this study was to identify indicators that could assist assessors in the identification of specific concerns earlier in the home study process.

The use of predicative analytics is a relatively new tool being used in child welfare systems in order to assist with decision making tools (Capatosto, 2017). In line with the emerging possibilities, indicators could statistically predict further compliance with licensing rules and systems and do not have a direct connection to risk levels. In other words, they are not the components of a home study that will be the basis of a home study denial or license revocation but may identify the need to look deeper into individual sections. The goal of identifying and incorporating predictive indicators within the home study is to improve decision-making and to support clinical judgment thereby increasing consistency and effectiveness of the tool itself. Furthermore, when an indicator is not found during an inspection it could predict further acceptable findings and could potentially shorten the amount of time spent on the home study itself saving (or reallocating) assessor time. For this purpose, there were two main research questions: (1) What are the singular thematic items within a home study based on the provider type and the licensing actions that inform the need for further investigation, (2) what are the thematic items within a home study that inform the need for further investigation based on frequency or patterns in relation to home study denials and revocations?

1.1. Definitions

- **Licensed Adoptive:** A provider that is licensed, or becoming licensed for the purpose of adoption.
- **Licensed General:** A provider that is licensed and is not identified as a relative or suitable other and is not currently planning to provide permanency (adoption or guardianship).
- **Licensed Kinship¹:** A provider that is licensed and is identified as being a relative to the foster child.
- **Licensed Suitable Other²:** A provider that is licensed and is identified as having a relationship with the foster child or the child's family prior to placement.
- **Unlicensed Adoptive:** A family member or suitable other that is seeking the home study for the purpose of adoption only and is not seeking licensure.
- **Unlicensed Kinship/Suitable other:** A family member or other that had a relationship with the child prior to placement and wishes to remain unlicensed.
- **Unified Home Study:** A comprehensive assessment that evaluates potential and/or current caregiver(s), and the physical environment and includes a recommendation pertaining to placement and permanency.
- **ICPC:** Interstate Compact on Placement of Children, refer to RCW 26.34

2. Methods and materials

A mixed method correlational exploratory research design was used to identify indicators (Greene, Caracelli, & Graham, 1989). This method was chosen to explore correlations of thematic identifiers within a sample of archived home studies. Through the review of individual home studies, qualitative descriptive themes were identified, both positive and potentially alarming, in order to map common concerns between the various archived home studies and their outcomes. Descriptors were coded into alphabetical (and ultimately numerical) themes allowing for quantitative analysis.

2.1. Participants

All Licensing Division foster care supervisors and area administrators, state administrators and Continuous Quality Improvement (CQI) personnel reviewed a total of 207 home studies. Supervisors from the 15 licensing offices (inclusive of satellite offices) not only participated but also identified one or two assessor(s) from each office to complete the reviews. In total, there were 15 supervisors, 15 licensors, six area administrators, three statewide administrators and three Continuous Quality Improvement (CQI) personnel who reviewed the home studies. Home study reviews were assigned across regions to ensure a limitation of personal bias by any supervisor, area administrator or licensor who may have personal investment or knowledge of any particular case. In addition, participants were asked to request a reassignment for any home study they may have personally worked on. In total, three cases were re-assigned during the process. Each participant reviewed approximately five home studies over a period of two months.

2.2. Sample

Two-hundred and seven (207) case studies were identified through a combination of convenient and proportionate stratified sampling

¹ Relative: As defined in RCW 74.15.020(2)(a), a person who is related to the child, expectant mother, or person with developmental disability...

² Suitable other: As defined in RCW 13.34.130, a person who the child or family has a preexisting relationship with...

Table 1
Total Percentage of Licensed State Home VS. Case Reviews per Region.

	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6	Total
Total Licensed Homes	746	390	498	392	381	857	3264
Total Case Reviews	49	23	31	24	24	56	207
Percentage	23%	11%	15%	12%	12%	27%	100%

measures ensuring statewide representation. The convenient sample included all home studies that were not recommended for approval (denial) and did not result in a foster care license as well as those that were approved but the license resulted in a revocation, or the cancellation of the license, any time after licensure. Qualifiers for the sample included: (1) Home studies had to be no older than 2013 because prior to that they weren't scanned into the electronic database, (2) Verification that the home study was available in the electronic database and, (3) Home studies had to be un-restricted (accessible to DCYF staff with normal permissions).

Secondly, all providers (licensed or unlicensed) that had an adoption that wasn't completed for any reason, and met the criteria listed above, were added to the convenient sample. The remaining amount of providers needed within regions included approved providers, both licensed and unlicensed, and were selected through proportionate stratified random sampling, assessed for usability and added to the sample ensuring the amount of cases mirrored the percentages of caseloads assigned to each of the state's six regions. [Table 1](#) represents the sample size according to the six regional caseload percentages.

As an added measure to ensure representation of the sample, the license type, or purpose, was also recorded. While the majority of home studies completed in this time period were considered "Unified" (completed for permanency, to include adoption) the identified purpose of the home study by the applicant was often for different purposes. This included; licensed adoptive, licensed general, licensed kinship, unlicensed adoptive, unlicensed kinship, unlicensed ICPC. [Table 2](#) outlines the home study type.

2.3. Tools

A data collection tool, produced in Microsoft Publisher and tested by the Continuous Quality Improvement (CQI) team was provided (in conjunction with detailed instructions) to each of the review teams. Data collectors were instructed to record all individual items listed in a home study, any recorded licensing violations including: intakes and/or findings pertaining to licensing infractions and/or allegations of child abuse, negligent treatment or maltreatment, and any documented concerns outside of licensing violations. For example, an unlicensed home study would not have licensing violations documented, but may include concerns that pertain to safety, well-being and/or permanency. Data collectors were also instructed to record the provider's statement of purpose within the home study (i.e. licensed, unlicensed, kinship, adoption, etc.). Once returned, data from each data collector was input with the corresponding ID number into the Excel tracking sheet specifically designed to identify licensing decisions, final requirements cited in the decision making process, and all individual items documented in the home study leading to the final recommendations. Coding was completed using a combination of Excel for thematic counts and the secondary analysis as well as IBM/SPSS version 26 for statistical analysis in order to determine the modifying effects of the overall demographics as well as identifying the predictors.

Table 2
Sample Provider Types.

Provider Type	Licensed Adoptive	Licensed General	Licensed Kinship	Unlicensed Adoptive	Unlicensed Kinship	Unlicensed ICPC
Count	39	105	14	12	36	1

2.4. Coding/analytcs

In order to create a data set for statistical analysis, it was critical to code qualitative data found within each home study. Each home study was de-identified and assigned a participation number in Excel and all themes were recorded for each case. Once the qualitative data was input into a spreadsheet, each statement was assigned a parent theme (or general classification based on the topic of each theme). Twenty-six parent themes (those qualitative topics that had general likeness) were identified and assigned followed by assignment of child themes (those items within a parent theme with detailed likeness based on the theme). For example, medical needs (MN) was identified as a parent theme and included specific conditions (child themes) such as diabetes, heart conditions, physical limitations and so on. 341 child themes were created to break the parent theme data into specific categories.

Once parent and child themes were assigned, the combination of all adult and child themes resulted in 64 final codes. These codes were then counted individually in each home study for the number of times they were mentioned. Through the coding process it was possible for themes to be mentioned more than once in a single home study. For example, an applicant may have expressed multiple medical needs (MN) such as issues with diabetes and a heart condition resulting in the code mentioned twice in one home study. Another example could be that an applicant mentioned they occasionally drink alcohol, smoke regularly and use marijuana resulting in the code drug and alcohol abuse or use (DAAU) being mentioned three times. The final code legend can be found in [Appendix A](#). Once the coding was completed, statistical analysis was applied.

2.5. Methodology and data analysis plan for indicators

The methodology used to generate the indicators was drawn and modified from the regulatory compliance and licensing research literature where this methodology has been used a great deal in making monitoring decisions (Licensing Key Indicator Methodology – LKIM (Fiene, 1985). The Licensing Key Indicator Methodology (LKIM) was used because it has proven to be very effective in dealing with nominally measured data that are extremely skewed (Fiene & Nixon, 1985).

The first step in the LKIM is to sort the frequency data from the various themes into high and low groups. This created a dichotomization of the frequency distribution that could then be used in a 2×2 matrix where each respective theme was compared to determine if it were in the high or low group (see [Chart 1](#) below). In other words, a Likelihood Ratio was calculated. Only those themes that either reached or exceeded a 0.75 agreement between having been observed and having the ability to distinguish between those homes where this was generally the case or not in the aggregate (total number of themes present). [Chart 1](#) provides a depiction of the relationship between each of the themes and the overall possible aggregate score of all the themes.

	High Group – Other Themes present	Low Group – Other Themes not present
Observed	Present	Absent
Not Observed	Absent	Present

Chart 1. Comparison of Individual Themes with High vs Low Group (Aggregated Themes).

2.6. Primary statistical analysis

IBM/SPSS Version 26 was used to analyze the data for this study. Basic demographics, (provider type and recommendation) and frequencies were run as well as a correlational analysis to explore relationships amongst the coded themes, licensing actions (inclusive of licensed revocations and unlicensed home study denials), and licensing violations. Third, analysis of variance (ANOVA) was done using one test to explore differences between provider types and the outcome variable of licensing violation data. And lastly, a 2×2 matrix was constructed in order to determine if specific codes could predict overall thematic success or failures by occurrence frequencies in either a high or low grouping. These analyses were performed on approximately 50 coded themes identified earlier in this paper and the various categories of homes: Licensed homes, Unlicensed homes, Licensing status of homes.

2.7. Secondary statistical analysis

Through the coding process it was possible for themes to be mentioned more than once in a single home study. For example, an applicant may have expressed multiple medical needs (MN) such as issues with diabetes and a heart condition resulting in the code mentioned twice in one home study. Another example could be that an applicant mentioned they occasionally drink alcohol, smoke regularly and use marijuana resulting in the code drug and alcohol abuse or use (DAAU) being mentioned three times. A secondary set of analytics was completed in order to identify codes based on frequency or multiple patterns within the same theme of a home study assessment.

This set of analytics was run in order to address how some items are not indicators until a pattern is discovered. This was done by first isolating all rows of the excel dataset indicating a denial (D) or revocation (R) regardless of provider type. Isolated items were placed into a 5 by 5 grid showing each combination of prevalent items. An example can be seen in [Chart 2](#). A divisive approach using complete linkage was then used taking the occurrence of the first matrix and applying it to the primary codes for comparison. This was done to identify the individual activities most commonly found in conjuncture with the codes within the denial and revocation status. Once both analyses were complete, significant codes found in the primary and secondary analyses were cross referenced to narrow the results to one overall set of indicators.

3. Results

It was found that the average number of licensing violations per home was just less than two (1.85 licensing violations), that the average length of a license was 41 months with a range of one to 137 months, and the average number of themed codes mentioned in a home study was eight. Correlational analyses determined there were significant

relationships between licensing violations and the licensing status of the homes, $r = -0.66$, $p < .0001$, $n = 207$.

Analysis of Variance (ANOVA) assessed the licensing violations by provider type and a significant difference was found ($F = 3.501$, $p < .005$, $n = 207$) with average violations ranging from 2.45 for licensed general homes to 2.10 licensed adoptive homes to 1.86 for licensed kinship homes. The following results (see [Fig. 1](#)) describes the codes found within the 2×2 primary analysis ([Chart 1](#)) as described in the methodology section above in which specific themes predict that other themes will be present in the aggregate. The results are presented with all the types of homes used, licensed homes, unlicensed homes, revocations, denials, compliant and finally licensed general homes. Note, the plus and minus within [Fig. 1](#) is visual representation that an indicator may or may not be positive or negative. For example, the experience of childhood trauma may be considered positive in a home study if it contributed to a person's resiliency and/or ability to relate to children who have experienced trauma. However, it may be considered negative if the childhood trauma is being re-experienced by the caregiver and it impacts their ability to provide care to children.

From [Fig. 1](#), it can be determined that educational success (ES) and positive family relationships (PFR) are the two coded themes that appear as themes in all the above listed categories. Home/Community Safety (HCS) was found in all areas except unlicensed homes: However, because HCS was found significant within denials, which is inclusive of unlicensed providers, this is an issue in unlicensed homes as well. Lack of stability of finances and work (FWUS) along with marriage/partnership issues (MPI) appear as other predictor indicators in four of the categories making them applicable to all provider types.

Logistic regression was then used by pairing the various codes within denial and revocation cases in combinations until there was no observable grouping which occurred in more than 24% of the time. A duplicated regression was done on cases where the same codes were present but the case did not result in a revocation or denial with the same results. This indicated that no grouping of two violations together were statistically relevant when there was a case of denial or revocation. Additionally, while looking at the home studies where single codes were found multiple times, we isolated the denial and revocation cases to remove any findings in non-problematic cases to remove false positives. The purpose of viewing the data through this particular lens was to remove any code that may have been found with multiple findings in cases that were never denied or revoked taken. The fact that some compliant rated home studies also showed some of these codes in duplicate did not disprove the finding's relevance as they are seen as possible predictive indicators that something may be going on that is critical in nature and warrants further investigation. If multiple findings of the same code were found in all three home study results then the number of times it was found in compliant cases and in denial or revocation were compared and only items that happened more in denial

CASES	AC	ANDA	ANDC	AUL	BCCL
AC	2	2	3	0	2
ANDA	4	9	4	1	3
ANDC	3	2	4	0	5
AUL	1	2	1	12	3
BCCL	1	3	0	3	18

Chart 2. 5 by 5 Secondary Cluster Analysis Example.

	CCI	AC	ES	FWUS	HCS	CTR	MPI	MH	RA	PFR	BCC	MRS	RLA	MI	UF	REN	FWS
1			+	-	+		-		+	+							
2			+	-	+					+	+						
3			+				-		+	+						-	
4	-	+	+	-	+	+-	-	+-	+	+	+	+	+				
5			+	-	+					+				+	+-		
6			+	-	+		-		+	+							
7			+		+		-			+	+	+					+

Fig. 1. Coded Themes and Categories of Homes. *Legend:* 1 = All Homes; 2 = All Licensed Homes; 3 = All Unlicensed Homes; 4 = Revocations; 5 = Denials; 6 = All Compliant; 7 = Licensed General Homes. CCI = CHILDHOOD CHALLENGES OR INSTABILITY; AC = APPLICANT COOPERATIVE; ES = EDUCATIONAL SUCCESS; FWUS = FINANCIAL/WORK UNSTABLE; HCS = HOME/COMMUNITY SAFETY; CTR = CHILD TRAIT REQUESTS; MPI = MARRIAGE/PARTNERSHIP ISSUES; MH = MENTAL HEALTH; RA = RESOURCES AVAILABLE; PFR = POSITIVE FAMILY RELATIONSHIPS; BCC = BACKGROUND CHECKS CLEARED; MRS = MARRIAGE RELATIONSHIP STABLE; RLA = RELIGIOUS AFFILIATION; MI = MILITARY; UF = UNSUPPORTIVE REFERENCES; REN = RELIGION NONE; FWS = FINANCIAL WORK STABLE.

and revocation cases were found.

In the 16 identified secondary indicators, less than 20% of compliant home studies showed multiple secondary indicators which was a factor used in refining all duplicate indicator findings down to the relevant ones for denial and revocation. This secondary cluster analysis of these repetitive codes indicated an overwhelming association that specific codes, considered minor in their own, may lead to denial or revocation when they are identified more frequently within the same home study. When sixteen of the codes used in identifying themes occurred more than once within the same case denial or revocation was recommended. It is important to note that not all violations appeared on each case, and not all had the same multiple findings of a specific code. Overall, home studies with more than one finding of these sixteen codes were denied or revoked in 76.6% of all cases in this study as demonstrated in Fig. 2.

The total 27 codes from the primary and secondary analysis were cross-referenced with one another in order to identify duplicate codes found in both analyses. Results are presented in Fig. 3. Because many of the codes were inherently the inverse of one another (i.e. Marriage/Partnership Instability vs. Marriage/Partnership Stability), each code was classified into three categories on how they were related between the two analyses; direct, indirect and not related. "Direct" means that the two codes found within both analyses were the same. "Indirect" means that while the codes were different, the theme was within the same intent of the parent theme. "Not related" references codes that were unique to the primary or secondary analysis. The purpose was to limit duplicative themes through highlighting bivariate associations.

As can be seen in Fig. 2, six codes from the secondary analysis emerged that were not related, either directly or indirectly, to the primary indicators. They include: abuse, neglect, domestic violence as an adult (ANDA), drug and alcohol abuse or use (DAAU), inappropriate discipline (DI), concerns identified with the family of origin (FOCI), medical needs (MN) and, self-identified challenges/lack of insight (SICLI).

4. Discussion

It is important to provide this preliminary study to introduce the parameters of locating and identifying emergent themes within the home study that could be used to guide depth of assessment as well as possible frequency of support. Through this study, we are able to identify 13 preliminary indicators within the foster care home study currently being used throughout the state. These categories combine both the inclusion and absence of the topic (i.e. educational success and educational challenges). They include; childhood experiences, level of cooperation, educational success, financial/work stability, home and/or community safety, child trait requested, marriage/partnership stability, mental health, availability of resources, current family relationships, background check clearance, religious affiliations, military, references.

Not all indicators within the first analysis applied to all provider types. Due to some codes being more significant within the varied provider types, this study demonstrates a pathway to differentiate a home study based on provider types (licensed versus unlicensed) and outcomes (revoked, denied and good standing). There are some indicators that apply only when considering a licensed general foster home study while others apply to a greater extent in unlicensed home studies. For example, marriage/relationship stable (MRS), background check clearance (BCC), applicant cooperative (AC), and childhood challenges or instability (CCI) only appeared as indicators in licensed general home studies while unsupportive references (UR), lack of religious affiliation (REN) and military (MI) only showed as indicators in unlicensed home studies. Likewise, several indicators are already specific for the type of home study needed. For example, child trait requested (CTR) is only applicable in licensed foster care because kinship care involves an identified family member or friend of the family, making the need to specify age, gender and ability of a child irrelevant.

Finally, the secondary analytics was able to identify items that while not an indicator when mentioned only once within a home study did become an indicator when found more than once. These repeated items

ANDA	ANDC	AUL	BCCL	DAAU	DFR	DI	ES	FOCI	FWUS	HCS	MH	MN	MPI	PFR	SICLI
0.91%	3.64%	7.27%	9.09%	0.91%	0.91%	1.82%	0.91%	0.91%	2.73%	4.55%	3.64%	0.91%	0.91%	0.91%	3.64%

Fig. 2. Percentage of codes mentioned more than Once in Denied or Revoked Home Studies. *Legend:* ANDA = ABUSE NEGLECT DOMESTIC VIOLENCE ADULT; ANDC = ABUSE NEGLECT DOMESTIC VIOLENCE CHILD; AUL = APPLICANT UNCOOPERATIVE/LIED; BCCL = BACKGROUND CHECK CLEARED; DAAU = DRUG ALCOHOL ABUSE/USE; DFR = DIFFICULT FAMILY RELATIONSHIPS; DI = DISCIPLIN INAPPROPRIATE; ES = EDUCATIONAL SUCCESS; FOCI = FAMILY OF ORIGIN CONCERNS IDENTIFIED; FWUS = FINANCIAL/WORK UNSTABLE; HCS = HOME/COMMUNITY SAFETY; MH = MENTAL HEALTH; MN = MEDICAL NEEDS; MPI = MARRIAGE/PARTNERSHIP ISSUES; PFR = POSITIVE FAMILY RELATIONSHIPS; SICLI = SELF IDENTIFIED CHALLENGES/LACK OF INSIGHT.

Indicator based on single occurrence	Indicator based on pattern or repeat occurrences	Indirectly related	Directly related	Not related
CCI	ANDC	X		
AC	AUL	X		
ES	ES		X	
FWUS/FWS	FWUS		X	
HCS	HCS		X	
CTR				X
MPI	MPI		X	
MH	MH		X	
RA				X
PFR	DFR		X	
BCC	BCCL	X		
MRS/MPI	MPI	X		
RLA				X
MI				X
UF				X
REN				X
FWS		X		
DFR		X		
ANDC		X		
AUL		X		
BCCL		X		
	DAAU			X
	ANDA			X
	DI			X
	FOCI			X
	MN			X
	SICLI			X

Fig. 3. Primary and Secondary Comparison: Identifying Duplicative and Dependent Findings. *Legend:* CCI = CHILDHOOD CHALLENGES OR INSTABILITY; AC = APPLICANT COOPERATIVE; ES = EDUCATIONAL SUCCESS; FWUS = FINANCIAL/WORK UNSTABLE; HCS = HOME/COMMUNITY SAFETY; CTR = CHILD TRAIT REQUESTS; MPI = MARRIAGE/PARTNERSHIP ISSUES; MH = MENTAL HEALTH; RA = RESOURCES AVAILABLE; PFR = POSITIVE FAMILY RELATIONSHIPS; BCC = BACKGROUND CHECKS CLEARED; MRS = MARRIAGE RELATIONSHIP STABLE; RLA = RELIGIOUS AFFILIATION; MI = MILITARY; UF = UNSUPPORTIVE REFERENCES; REN = RELIGION NONE; FWS = FINANCIAL WORK STABLE; ANDA = ABUSE NEGLECT DOMESTIC VIOLENCE ADULT; AUL = APPLICANT UNCOOPERATIVE/LIED; BACKGROUND CHECK CLEARED; DAAU = DRUG ALCOHOL ABUSE/USE; DFR = DIFFICULT FAMILY RELATIONSHIPS; DI = DISCIPLIN INAPPROPRIATE; FOCI = FAMILY OF ORIGIN CONCERNS IDENTIFIED; MN = MEDICAL NEEDS; SICLI = SELF IDENTIFIED CHALLENGES/LACK OF INSIGHT.

began to show patterns within certain codes. The cases in which they did happen showed a significant chance of having other issues which were considered problematic. This leads to the conclusion that items found in groupings also meets the indicator definition and shows that minor findings in key patterns may create scenarios that, while not being serious enough to cause a denial or revocation themselves, warrant further investigation and or mitigation strategies before a home study is approved. For example, an assessor may not have enough information to justify a denial and also have concerns that are preventing them from approving the home study, but upon finding multiple sightings of the same minor indicators it could give reason for further investigation in some areas.

Four specific limitations were identified; sample size, consistency of data collection, individual bias and systematic bias. Each of these limitations were known from the beginning of the project and while addressed throughout could not be completely mitigated.

Because home study data points are not recorded in an electronic database, consideration to the amount of time and limited resources available was considered. There are approximately 2250 homes studies

completed each year; this sample (207) represents only 9.2% of the annual statewide home study caseload. Ideally, the sample would be 20% of all home studies totaling closer to 400 home study reviews in order to provide smaller margins for error and increase the capacity of finding significant differences within the codes.

While data collection training was provided, the fact that all 207 home studies were reviewed by 39 data collectors with varying degrees of effort and interpretation resulted in unknown and potentially limited inter-rater reliability. This can be seen by the fact that recorded themes on the homes studies varied between one thematic discovery to 48 discoveries. In order to address this limitation, resources need to be provided to limit data collector numbers as well as provide precision training.

Another challenge identified was individual assessor bias. The fact that home studies are summary assessments with very broad categories inherently lends to predispositions or individual bias. While assessors are provided guidance and general questions to ask the applicants, the results are often open to interpretation and may be influenced by individual bias of the assessor themselves. Due to the racial, ethnic,

cultural, gender, and economical predispositions of the home study team, unintended bias within the home study could have led to a distortion in the original data. The level of impact of individual bias to this study is unknown.

Finally, DCYF is committed to identifying and limiting systematic racial disparity within the agency's entire child welfare system. In a 2018 report, it was found that the rate of children entering placement in Washington varied depending on the ethnic group. For example, children of native American Indian/Alaskan Native (AI/AN) were removed from their homes at a higher proportional rate than children of other ethnic groups. Placement of black children/youth were only slightly higher than white children and children of Asian or Pacific Island ethnicity were less likely to be placed than white children (Graham, 2019). In contrast, the racial and ethnic make-up of the foster care provider community in Washington State is largely white at approximately 75%. Because of this known limitation, it is possible disproportionality and systemic bias will continue to be reflected through the use of the identified indicators.

5. Recommendations

We assume the use of indicators may have the potential to reduce subjective decision making and bias by creating an avenue for consistency in home study assessments and guidance for recommended approvals or denials. This can be done by identifying areas that may predict unsuccessful placement. It is possible this will provide guidance to assessors when further assessment and mitigation is needed. Additionally, with a carefully designed home study tool it should be possible to reduce the amount of time an assessor uses to approve or deny a home study, saving agency time and resources. Finally, by differentiating and focusing technical assistance to those applicants that need more or specific support, the use of indicators may increase the success of timely placement and permanency. However, in order to assess and evaluate if the above assumptions are correct, it is critical to design a system that collects the data for validation and reliability purposes.

In order to use the indicators, they must be systematically and intentionally imbedded in the home study sections as direct questions which could be pre-populated based on the applicant's responses on the application. Due to the potential bias limitations, indicators should not be used as a tool to recommend or not recommend placement or permanency. Therefore, it is recommended that indicators are used as a tool to identify potential areas that need additional information within a home study and require mitigation assessments as needed. There are two recommendations for indicator use: 1) further assessment before determination and, 2) frequency of monitoring post assessment.

Further Assessment: Indicators, when encountered within a home study, could require an assessor to gather further information before assurances of mitigation can be determined. For example, if it is determined one or more applicant did not complete a high school (ES) degree or GED an assessor would ask additional questions pre-identified by the system such as "What were the factors that contributed to the incomplete education?" and "What are the belief systems regarding education for potential foster children?". If an assessor determines a mitigation is successful based on the interview responses, they move onto the next section. If an assessor determines a mitigation is not successful, or there is not enough information, they could require additional actions such as training or the review of a topic-specific resource.

Frequency of Monitoring: A second strategy that could be used with

indicators is to set a threshold of key indicators within any one home study whereby additional, or more frequent post licensing monitoring visits should be completed. This strategy would only be applicable for licensed providers as unlicensed providers are not required to have ongoing monitoring. For example, if three different indicators are present determined by the assessor's review of the applicants' answers, the recommendation could be made that a health and safety check will be completed within six months' time, upon placement. If four or more indicators are present, a health and safety check could be required with 4 months of placement, etc.

It is important to continuously evaluate the predicative indicators within the home study assessment. This is to ensure comprehensive and up-to-date data are considered, how changes to policy and practice affect the model, and whether the model needs to be changed or modified. Ongoing evaluation should consider if staff are implementing and using the home study approach as intended. Validation should include a larger data set of home study components inclusive of these identified indicators.

6. Conclusion

There are many potential uses of indicators within the home study process. Once Washington systemizes how they will be used, additional validation studies will be developed to ensure appropriate identification and use as well as identifying any unintended consequences before moving forward with implementation. It is critical to remember that thematic identification of indicators does not definitively inform the assessor of the outcome or determine the recommendation of the home study. However, it is possible indicators may help improve consistency by identifying emerging patterns and limiting variations in decision-making. They could also help ensure applicants statewide are being assessed and treated similarly.

The purpose of the home study is to assess the caregiver(s)' ability to provide a safe home, the quality of care needed by children and an environment that is nurturing, respectful and supportive (WAC 110-148-1320, 110-148-1365). To measure whether or not that goal is being met, ongoing research and validation is needed. This study was conducted to initially identify indicators or predictors that would assist licensors with identification of broader concerns early on in the home study. This study classifies what are considered universal, licensed (general), unlicensed (kinship) and repeat or pattern indicators. All of which will help with identifying what areas of the home study should have more, or less, depth.

There was no additional funding secured for this study

CRediT authorship contribution statement

Sonya Stevens: Conceptualization, Methodology, Data curation, Project administration, Supervision, Writing - original draft, Resources.
Richard Fiene: Formal analysis, Validation, Writing - review & editing.
Daniel Blevins: Conceptualization, Formal analysis, Validation.
Amber Salzer: Resources, Data curation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Code identification sheet

Code	Theme	Description
ACAC	Adult child Adverse Experiences	One or more adult child(ren) of an applicant experienced adverse experiences such as abuse, homelessness, substance use or mental health issues.
ANDC	Abuse Neglect Domestic Violence Child	One or more applicant experiences Abuse and/or neglect as a child which may include inappropriate discipline such as spanking, restraining or withholding food.
CCI	Childhood Challenges or Instability	Applicant reported non-abuse related challenges during childhood such as (but not limited to) instability, bullying at school or difficulty with friends.
BCCL	Background Check Issues Identified (including license actions)	One or more applicant has one or more issues identified during a background check which may have included one or more of the following: Criminal, founded findings, licensing violations, traffic violations, etc.
HB	Healthy Boundaries	The applicant has healthy relationships and boundaries with others
AC	Applicant cooperative	The applicant was cooperative during the licensing process
AUL	Applicant Uncooperative or Lied	One or more applicant was either uncooperative during the licensing process by delaying or refusing to meet requirements or lied or withheld information during interviews or on paperwork.
DA	Discipline appropriate	Discipline and guidance used in the home is appropriate
DI	Discipline Inappropriate	Discipline and/or guidance used in the home is inappropriate including but not limited to spanking, isolation, food restrictions, etc.
ID	Inclusive of Diversity	Inclusive of Diversity including race, ethnicity, ability and sexual orientation
DD	Difficulty Diversity	Difficulty with Diversity including but not limited to race, ethnicity, ability and sexual orientation
ECDO	Educational Challenges or Drop out	One or more applicant experienced challenges in school such as dropping out, being expelled or self-inflicting unrealistic expectations toward education.
ES	Educational Success	One or more applicant was successful in school through high school or GED completion and perhaps completing some or all of a college degree.
FWS	Financial/Work Stable	The applicants reported financial stability as well as having stable employment.
FWUS	Financial/Work Unstable	The applicant reported unstable personal finances and/or unstable work meaning unemployment, frequent changes in employment or inability to work.
HCS	Home/Community Safety	The licensor reported one or more concerns regarding the safety of the home or community such as mold, unstable stairs or floors, elements in or near the home such as bodies of water, highways, or unsafe neighborhoods.
HM	Home Moves	Relocation (frequent or recent)
RM	Roommate	Roommate in the home
CTR	Child Trait Requests	Specific child trait restrictions or requests for placement
FA	Foster/Adoption	Prior foster/adoption
MPI	Marriage/Partnership Issues	One or more applicant reported instability or difficulties in their marriage, domestic partnership or long term relationship.
MC	Medically cleared	The applicant was medically cleared by their physician
MN	Medical Need	Specific medical needs or areas of concern were identified by the physician
MH	Mental Health	Significant mental health issues
DAAU	Drug/Alcohol Abuse/Use	One or more applicant reported using drugs and alcohol either and don't do it anymore or only drink occasionally.
RA	Resources Available	There are resources the applicant reported as available
SICLI	Self-Identified Challenges Lack of Insight	Either the licensor or one or more applicant reported self-behavior challenges such as but not limited to the inability to know how prior experiences affect current behavior or an understanding of parenting.
WEPL	Work/experience parenting limited	One or more applicant reported having limited parenting experience
FP	Family Planning	The applicant reported fostering or adoption as a means to grow the family
FCCSN	Family of creation child(ren) special needs	The applicant has one or more children with disabilities
FC	Foster child(ren)	The applicant has a foster child in the home
OA	Origin Adoption	The family of origin experienced adoption
MI	Military	An applicant has military experience
FOPAR	Family of Origin Parental Relationships	One or more applicant reported their own parents struggled in marital or partnership relationships such as divorce or multiple marriages
UF	Unsupportive References (including children in home)	References were limited or do not support the applicant including formal references and comments made by children in the home and/or adult children.
PFR	Positive Family Relationships	One or more applicant reported positive family relationships including those in both the family of origin and family of creation.
DFR	Difficult Family Relationships	One or more applicant reported difficult relationship(s) between themselves and the family of creation or origin.
FOCI	Family of Origin Concerns Identified	One or more of the applicants reported serious concerns with someone in their family of origin including but not limited to addiction, criminal record or health related issues.
ANDA	Abuse Neglect Domestic Violence Adult	One or more applicant reported experiencing abuse and neglect as an adult including but not limited to domestic violence, physical abuse, emotional abuse. This experience may have been to be a witness.
RU	Resources Unavailable	There are a lack of resources available to the applicant
HG	Home Good	The licensor noted the home is physically good, safe and stable in general
BCC	Background Check Cleared	No issues were identified about one or more applicant during the background check process.
BSP	Blended Shared Parenting	Within the home there are children of creation from multiple relationships usually inclusive of shared parenting with a parent outside the home.
CEPG	Childhood Experiences Positive General	One or more applicant reported generally positive childhood experiences
EHS	Education Home School	The applicant home schools their children.
MRS	Marriage Relationship Stable	The applicants reported a stable marriage, domestic partnership or long term relationship.
REN	Religion None	One or more applicant reported they did not identify as religious or spiritual.
RLA	Religious Affiliation	One or more applicant reported an affiliation to a religion or spiritual community.
RE	Related Experience	One or more applicant reported experience they believe would help them to be a foster parenting either through parenting or through work such as education, medical, civil service, parenting other children, parenting classes, etc.
BRUG	Boundaries/relationship unhealthy general	The licensor reported unhealthy relationships between the applicant and others in general.
VGA	Values/Goal/Attitude Positive	One or more applicant reported strong insight and/or positive goals and values.

Appendix B. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.childyouth.2020.105133>.

References

- Capatosto, K. (2017). Foretelling the future: A critical perspective on the use of predictive analytics in child welfare. Kirwan Institute for the Study of Race and Ethnicity at The Ohio State University. Retrieved from <http://kirwaninstitute.osu.edu/wp-content/uploads/2017/05/ki-predictive-analytics.pdf>.
- Child Welfare Information Gateway (2018). Home study requirements for prospective foster parents. Retrieved from <https://www.childwelfare.gov/topics/systemwide/laws-policies/statutes/homestudyreqs/>.
- Conradi, L., Landsverk, J., & Wotring, J. R. (2014). Integrating safety, permanency and well-being series: Screening, assessing, monitoring outcomes and using evidence-based interventions to improve the well-being of children in child welfare. *Children's Bureau: An Office of Administration for Children and Families*. https://www.acf.hhs.gov/sites/default/files/cb/wp2_screening_assesing_monitoring.pdf.
- Crea, T. M., Barth, R. P., & Chintapalli, L. K. (2007). Home study methods for evaluating prospective resource families: History, current challenges, and promising approaches. *Child Welfare: Journal of Policy, Practice, and Program*, 86(2), 141–159.
- Cuccaro-Alamin, S., Foust, R., Vaithianathan, R., & Putnam-Hornstein, E. (2017). Risk assessment and decision making in child protective services: Predictive risk modeling in context. *Children and Youth Services Review*, 79, 291–298. <https://doi.org/10.1016/j.chldyouth.2017.06.027>.
- Depanfilis, & Girvin (2005). Investigating child maltreatment in out-of-home care: Barriers to effective decision-making. *Children and Youth Services Review*, 27, 353–374. <https://doi.org/10.1016/j.chldyouth.2004.11.010>.
- Department of Children, Youth, and Families. (2008). Policy 45274: Placement with unlicensed relatives or suitable persons. Retrieved from <https://www.dcyf.wa.gov/4527-kinship-care-searching-placing-and-supporting-relatives-and-suitable-other-persons/45274>.
- Fiene & Nixon (1985). Instrument based program monitoring and the indicator checklist for child care. *Child Care Quarterly*, 14(3), 198–214.
- Fiene (1985). Measuring the effectiveness of regulations. *New England Journal of Human Services*, 5(2), 38–39.
- Graham, J.C. (2019). Washington state DCYF racial disparity indices report (2018). Washington State Department of Children, Youth, and Families; Office of Innovation, Alignment, and Accountability. Retrieved from https://www.dcyf.wa.gov/sites/default/files/pdf/reports/Washington_State_DCYF_Racial_Disparity_Indices_Report_2018.pdf.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Educational evaluation and policy analysis. *American Educational Research Association*, 11(3), 255–274. <https://doi.org/10.2307/1163620>.
- Rossi, P. H., Schuerman, J., & Budde, S. (1999). Understanding decisions about child maltreatment. *Evaluation Review*, 23(6), 579–598. <https://doi.org/10.1177/0193841X9902300601>.
- Washington Administrative Code (2020). WAC 110-148-1320, 110-148-1365. Retrieved from <https://apps.leg.wa.gov/wac/>.

Improving Child Care Quality Through an Infant Caregiver Mentoring Project

Richard Fiene

The Pennsylvania State University

ABSTRACT: An evaluation of a mentoring training program for infant caregivers is described. Fifty-two infant caregivers from 27 childcare center-based programs were involved in a four month long intervention in which they were paired with an experienced early childhood educator. The focus of the mentoring program was to improve the overall quality of the classroom environment, as well as making the caregivers more sensitive to the needs of the infants. The results clearly indicated that the mentoring program was very effective in improving the overall quality of the classroom, as well as making caregivers more sensitive to infants' needs.

KEY WORDS: infant caregivers; childcare; mentoring; training.

Introduction

This paper describes a child care mentoring project designed to improve the quality of infant and toddler child care programs in south central Pennsylvania. The goal of the mentoring project was to improve the quality of the child care environment and specifically the quality of caregiver-child interactions. As most caregivers in Pennsylvania only receive workshop training, the goal of this project was to compare the mentoring approach to the more typical workshop training. Mentoring is being explored because of its targeted intensive one-on-one nature in delivering training to caregivers based upon needs assessments. The project was conducted during the later half of 2000 and the beginning of 2001. The results presented in this paper are part of the pre- and post-test data collection phase (summer 2000 and winter 2000–2001) of this mentoring project. The actual mentoring intervention occurred from September through December 2000.

Correspondence should be directed to Richard Fiene, Ph.D., Director, Capital Area Early Childhood Training Institute, and Senior Research Associate, Prevention Research Center, The Pennsylvania State University, 2001 North Front Street, Suite 314, Harrisburg, PA 17102; e-mail: rjf8@psu.edu, Web page: <http://caecti.org>.

Mentoring in childcare has been documented in the literature for the past 10–15 years (Breunig & Bellm, 1996; Fenichel, 1992). It has been demonstrated to be an effective mode of training/technical assistance (Breunig & Bellm, 1996). However, in the majority of studies conducted there are few, if any, demonstrations that utilize a randomized trial design (Breunig & Bellm, 1996). Many studies track the progress of the intervention group, some studies have comparison groups, but few, if any, have employed a randomized design. This research paper will describe the pre- and post-test data collected as part of a study that has employed a randomized design.

The majority of research (Clarke-Stewart, 1987; Goelman & Pence, 1987; Howes, 1987; Phillips, 1987; Kontos & Fiene, 1987; Galinsky, Howes, Kontos, & Shinn, 1994; Scarr, Eisenberg, & Deater-Deckard, 1994; Iutovich, Fiene, Johnson, Koppel, & Langan, 1997; Helburn, 1995; Fiene, 1995, 1996; Jorde-Bloom, 1988; Love, Schochet & Meckstroth, 1986) completed on early childhood quality has focused on pre-school programs, with infant toddler programs rarely as the central focus of the research. The research completed in infant toddler programs has clearly documented the mediocre level of care provided to children in these programs (Iutovich, Fiene, Johnson, Koppel, & Langan, 1997). In the present study, we focus on the first three years of life. All the centers and the classrooms reported upon in this study serve children from birth to less than three years of age.

This report is organized as follows: a methodology section briefly describes the sample selected with basic demographic information on directors, caregivers and the programs. This is followed by a results section that provides pre- and post-test average scores for each of the assessment tools utilized in this study to measure quality, caregiver behaviors, knowledge, and organizational climate of programs. This section is followed with a discussion section and implications regarding this mentoring project.

Methods

Study Design

This study involved 52 caregivers from 27 sites in south central Pennsylvania. All programs were child care centers licensed by the Department of Public Welfare. Seven of the sites were accredited by the National Association for the Education of Young Children.

This study employed a randomized design in which a self-selected group of programs and caregivers were randomly assigned to two groups, either the mentoring group or the comparison non-mentoring

comparison/control group. Intervention model mentoring group received intensive mentoring from a seasoned early childhood professional (minimum of 5–7 years of experience in the early childhood field as both a director and teacher) from September to December 2000. The mentoring model consisted of a problem solving approach in which the mentor spent a good deal of time observing in the beginning weeks in order to develop a trusting relationship with the protégé. Once both the mentor and protégé felt comfortable then suggestions could be entertained by the mentor.

The comparison group did not receive the mentoring intervention and only had the regular workshop type variety training available to them. However, the comparison group did receive mentoring during the Spring 2001 from March to June 2001. What is of interest in this study is to determine how much the two groups have improved from the pre-test data collection because they were essentially equivalent at that point on all measures.

Programs were recruited by the Capital Area Early Childhood Training Institute, a broad based community focused training institute. Program directors were invited to attend a meeting describing the mentoring project. Of those attending, 95% agreed to participate in the project. Fifty two caregivers started the project, 14 caregivers dropped out of the project between pre- and post-test. There was an equal drop out rate from both the mentoring and the control groups.

Data from the four quality measures used for all the programs are presented in Table 1. The four measures of quality were the Infant Toddler Environment Rating Scale (ITERS), the Arnett Caregiver Observation Scale, the Knowledge of Infant Development (KIDI), and the Bloom Scales of Organization Climate.

The program directors' average age is 31 with a range from 24–53

Table 1
ITERS, Arnett, KIDI, Bloom Scale Scores

All Programs (n = 38)	Pre-Test	Post-Test	Change	Significance
ITERS	134	140	+6	ns
Arnett	30	40	+10	ns
KIDI	14	14	-0-	ns
Bloom	78	79	+1	ns

years of age. They are predominantly Caucasian (81%). Eight percent have associate degrees, 78% have bachelor's degrees, and 14% have master's degrees. They had been employed as directors in their program for an average of 31 months with a range from 1 month to 120 months. Their average pay is between \$20000–25000 per year. Sixty percent have health insurance and 45% have some form of dental or life insurance. Forty-five percent are in a retirement system.

The average age of caregivers in the programs was 36 with a range from 18–68. They are predominantly Caucasian (77%). Fifty-seven percent have high school diplomas, 16% have some college credits, 5% have CDA's, 16% have associate degrees, 5% have bachelor's degrees, and 2% have master's degrees. They have been employed as caregivers in their program for an average of 34 months with a range from 1 month to 153 months. They have worked in the early childhood field as caregivers for an average of 71 months with a range from 1 month to 312 months. Their average pay is between \$10000–15000 per year. Fifty percent have health insurance and 33% have some form of dental or life insurance. Thirty-three percent are in a retirement system.

The average size of the centers is 98 children with 17 staff employed either full time or part time at the program. The average weekly fee for infant care is \$137.00 per week and for toddler care is \$124.00 per week. The majority of staff are employed at the centers for either less than 1 year or greater than 5 years.

Results

Both the mentoring and comparison groups were tested for equivalence at the beginning of the project in the pre-test data collection phase. There were no statistically significant differences on any of these measures at the pre-test. When the programs and caregivers were measured at the post-test, positive changes occurred although none were found to be statistically significant. In the aggregate, the programs that continued with the mentoring project showed improvements in the overall quality of care.

Tables 2 through 5 present the pre- and post-test data for the intervention and control groups.

These results indicate that the mentoring group showed increases on the program quality scales (ITERS and Arnett). This increase is especially noticeable on the ITERS. Further, there was a decrease in program quality with the control group, going from a score of 137 to 132. On the Arnett scale the mentoring group increased greater than the control group (11 point increase versus a 7 point increase).

Although the above results did not reach statistical significance,

Table 2
ITERS

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	134	141	+7	ns
Control Group	137	132	-5	ns

Table 3
Arnett

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	29	40	+11	ns
Control Group	33	40	+7	ns

Table 4
KIDI

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	14	14	-0-	ns
Control Group	14	15	+1	ns

Table 5
Bloom

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	73	74	+1	ns
Control Group	87	91	+4	ns

when specific subscales are analyzed several show significant differences (see tables 6 and 7). Several of the subscales on the ITERS and Arnett reached statistical significance with positive changes in routines (greeting/departing, meals/snacks, nap time, diapering/toileting, health/safety practice/policy) learning activities (eye-hand coordination, active physical play, blocks, pretend play, cultural awareness), sensitivity, and appropriate discipline for the mentoring group. The only statistically significant finding with the control group was in a negative change in interactions in which the scores decreased from pre-test to post-test. Paired t-tests were used in all of these analyses for Tables 6 and 7.

Table 6
Mentoring Group

	Pre-Test	Post-Test	Significance
ITERS subscales			
Routines	36	41	.005
Listening activities	8	9	ns
Learning activities	28	31	.05
Interactions	13	13	ns
Adult needs	17	19	ns
Arnett subscales			
Sensitivity	26	31	.001
Appropriate discipline	7	9	.05

Table 7
Control Group

	Pre-Test	Post-Test	Significance
ITERS subscales			
Routines	41	42	ns
Listening activities	9	8	ns
Learning activities	29	31	ns
Interactions	15	13	.02
Adult needs	17	17	ns
Arnett subscales			
Sensitivity	28	31	ns
Appropriate discipline	6	7	ns

Discussion

These data demonstrate that the sites that were mentored improved on the ITERS and the Arnett. This is an encouraging result in that the intervention was only 4 months long. It is an important finding because the majority of mentoring projects in the past have utilized anecdotal evidence to demonstrate their effectiveness. Very few programs have conducted randomized trials of their interventions.

It is clear from the data that training/technical assistance interventions are needed in infant toddler programs because of the low scores on various program quality measures. It is also discouraging in that the control programs did not improve in which the ITERS went from 137 (pre-test) to 132 (post-test). This is a finding that will be monitored over time to see if this trend continues. Hopefully this was just an aberration in the data; however there does seem to be support when these data are compared to other studies (Iutovich, Fiene, Johnson, Koppel, & Langan, 1997).

The public policy implications are that an intensive mentoring intervention of only four months can produce positive, although not statistically significant, changes in the overall quality of child care programs both globally and with caregiver interactions. Previous research (Johnson, 1994) has indicated that increasing the number of hours of training produces more developmentally appropriate behaviors in child care staff. Mentoring fits this model because it is an intensive one on one intervention in which the mentor and protégé are engaged in problem

solving activities to improve the overall quality of the interactions and environment of the child care program.

References

- Behar, L. & Stringfield, S. (1974). A behavior rating scale for the preschool child. *Developmental Psychology*, 10, 601–610.
- Breunig, R. & Bellm, D. (1996). *Early childhood mentoring programs: A survey of community initiatives*. Washington, D.C.: Center for the Child Care Workforce.
- Clarke-Stewart, A. (1987). In search of consistencies in child care research. In D. Phillips (Ed.), *Quality in childcare: What does research tell us?* (pp. 105–119). Washington, DC: National Association for the Education of Young Children.
- Feagans, L. & Farran, D. (1979). *Adaptive Language Inventory*. Unpublished instrument. University of North Carolina, Chapel Hill.
- Fenichel, E. (1992). *Learning through supervision and mentorship*. Washington, D.C.: Zero to Three: National Center for Infants, Toddlers and Families.
- Fiene, R. (1995). Utilizing a statewide training system to improve childcare quality. *Child Welfare*, 74(6): 1189–1201.
- Fiene, R. (1996). Using a statistical-indicator methodology for accreditation. In S. Bredekamp (Ed.), *A Decade of Accreditation*. Washington, DC: National Association for the Education of Young Children.
- Galinsky, E., Howes, C., Kontos, S., & Shinn, A. (1994). *The study of children in family childcare and relative care: Highlights of findings*. New York, NY: Families and Work Institute.
- Goelman, H. & Pence, A. (1987). Effects of childcare, family, and individual characteristics on children's language development: The Victoria day care research project. In D. Phillips (Ed.), *Quality in childcare: What does research tell us?* (pp. 89–104). Washington, DC: National Association for the Education of Young Children.
- Griffin, A. & Fiene, R. (1995). *Childcare planning and evaluation*. Washington, DC: Zero to Three, National Center for Infants and Toddlers.
- Harms, T. & Clifford, R. (1980). *The early childhood environment rating scale*. New York: Columbia University Teacher's College Press.
- Helburn, S. (1995). *Cost, quality and child outcomes in child care centers*. Denver, CO: Center for Research in Economics and Social Policy, Department of Economics, University of Colorado.
- Howes, C. (1987). Quality indicators in infant and toddler childcare: The Los Angeles study. In D. Phillips (Ed.), *Quality in childcare: What does research tell us?* (pp. 81–88). Washington, DC: National Association for the Education of Young Children.
- Hresko, W., Reid, D. K., & Hammill, P. D. (1981). *The Test of Early Language Development*. Austin, TX: Pro-Ed.
- Iutovich, J., Fiene, R., Johnson, J., Koppel, R., & Langan, F. (1997). *Investing in our children's future*. Erie, PA: Keystone University Research Corporation.
- Johnson, J. (1994). *Child care training and developmentally appropriate beliefs and practices of child care employees in Pennsylvania*. Harrisburg, PA: Center for Schools and Communities.
- Jorde-Bloom, P. (1988). Assess the climate of your center: Use the early childhood work environment survey. *Day Care and Early Education*, Summer 1988: 9–11.
- Katz, L. (1994). Perspectives on the quality of early childhood programs. *Phi Delta Kappan*.
- Kontos, S. & Fiene, R. (1987). Childcare quality, compliance with regulations, and children's development: The Pennsylvania study. In D. Phillips (Ed.), *Quality in childcare: What does research tell us?* (pp. 57–79). Washington, DC: National Association for the Education of Young Children.
- Love, J., Schochet, S., & Meckstroth, R. (1986). *Are they in any real danger: What*

- research does—and doesn't—tell us about childcare quality and children's well being.* Princeton, NJ: Mathematica Policy Research, Inc.
- Maternal and Child Health Bureau (1997). *Stepping stones to Caring for Our Children*, Washington, D.C.: Maternal and Child Health Bureau, Department of Health and Human Services.
- Morgan, G., Costley, C., Genser, A., Goodman, M., Lombardi, J., & McGimsey, J. (1993). *Making a career of it: The states of the states report on career development in early care and education*. Boston: The Center for Career Development in Early Care and Education, Wheelock College.
- NAEYC (1987). *Accreditation standards*, Washington, D.C.: National Association for the Education of Young Children.
- Scarr, S., Eisenberg, E., & Deater-Deckard, R. (1994). Measurement of quality in child care centers. *Early Childhood Research Quarterly*, 9: 131–151.

Measurement Tools and Systems

Chapter 11

*Richard Fiene PhD
Karen E Kroh*



■ TABLE OF CONTENTS

Introduction	1
Definitions	3
Instrument Based Program Monitoring	5
Checklists	7
Rating Scales	9
Weighting Systems.....	11
Licensing Indicator Systems	19
Outcome Based Systems.....	27
Relationship Between Rules and Instruments	29
Reliability and Validity	31
Balance Between Compliance and Program Quality	33
Conclusion.....	35
References	37

■ PREFACE

The purpose of this chapter is to acquaint the licensing administrator with the science and art of measurement as it relates to regulatory administration. It is becoming more and more critical that licensing administrators have at least a rudimentary knowledge of measurement methods. Measurement is a key element of the new information age. It is the basis for the design and implementation of information systems, either manual or automated, conducting on-site inspections, making observations, interviewing and completing complaint investigations.

This chapter provides an overview to the major types of measurement tools used within the regulatory administration field related to assessing compliance with human care licensing rules. A historical perspective will be provided followed by outlines of key definitions. The types of measurement tools and systems will be reviewed. The final section of this chapter will address the relationship between measurement and rule formulation.

The sections titled Weighting Systems and Licensing Indicator Systems are heavily influenced by the two papers written by NARA Immediate Past President Karen E Kroh, Pennsylvania, in the late 1980s on these two topics.

Past NARA Secretary and Vice President Carolynne H Stevens, Virginia, and NARA Executive Director and Past President Pauline D Koch, Delaware, served as reviewers for this chapter.

■ INTRODUCTION

Measurement within regulatory administration has changed substantially from the 1970s through the 1990s. It has moved from being very qualitative to being more quantitative in nature. The qualitative nature was depicted with long narratives obtained from in-depth observations and interviews that described a facility in detail with a listing of violations with specific rules. The observations used a running record format in which a detailed accounting of the facility was obtained. This is in contrast to an anecdotal type of record that is used a great deal in the measurement literature related to observing behaviors. This qualitative system worked well when there were few facilities to be assessed. However, as the number of human care facilities increased and licensing agency administrators felt a greater need to understand compliance trends, movement to a more quantitative measurement system has evolved.

This move to quantification of measurement began in earnest in the 1970s, in particular, with the revision of the Federal Interagency Day Care Regulations (FIDCR). The notion of an instrument based program monitoring or licensing system began to be examined by licensing agencies. Checklists and rating scales were employed, with checklists being used predominantly because of the nature of regulatory compliance. However, a few states, provinces and cities utilized rating scales to measure compliance with rules. More will be said about the differences between checklists and rating scales.

By the early 1980s with severe federal cutbacks in funding, licensing administrators found themselves with an increasing number of facilities to license but fewer funds to perform the investigative function. In response to this concern, the indicator checklist methodology was created which utilized a shortened version of the comprehensive checklist approach used by many states. Indicator systems have been developing over the past two decades and in many states are key components of their monitoring and licensing functions. The indicator system is only one form of what is known in the licensing literature as inferential inspections. However, only the indicator system will be addressed in this chapter because the other types of inferential inspections are not valid and reliable enough to meet the criteria for scientifically based measurement tools.

■ INTRODUCTION

A related but very different technique that complements indicator systems is the use of weighting systems to determine the relative risk of specific rules related to non-compliance. The reason for the development of weighting systems is the nature of regulatory compliance data. Because compliance data measure minimum health, safety and well-being rules, the data are highly skewed with very little variance. The use of weighting systems helps to increase the amount of variance in the regulatory data sets.

The indicator and weighting systems have not been limited to licensing systems but have also been developed for other program quality endeavors such as accreditation and national standards setting.

A very recent development, in the 1990s, is the development and use of outcome based systems for licensing. This is where a licensing agency places more emphasis on outcomes rather than processes. This is a very experimental and controversial development, particularly for the field of human care licensing.

■ DEFINITIONS

Instrument Based Program Monitoring

A movement within licensing and regulatory administration from qualitative measurement to a very quantitative form of measurement that includes the use of checklists.

Indicator System

A licensing measurement system utilizing a shortened version of a comprehensive checklist measuring compliance with rules through a statistical methodology. Only key predictor rules are included on an indicator checklist. It is a form of inferential inspections where only a portion of the full set of rules is measured.

Inferential Inspections

An abbreviated inspection utilizing a select set of rules to be reviewed. An indicator system, weighting of rules for determining a shortened inspection tool, a random selection of rules, etc. are examples of inferential inspections. The use of inferential inspections by licensing agencies was developed as a time saving technique and a technique to focus regulatory efforts on facilities that required additional inspections or technical assistance.

Checklist

A simple measurement tool that measures compliance with state rules in a yes/no format. Either the facility is in compliance with rules or not in compliance. Generally, there is no partial compliance with checklists generally.

■ DEFINITIONS

Rating Scale

A more complex measurement tool in which a Likert type of rating is employed—going from more to less, or high to low. A rating scale is always used in the development of weighting systems. It is not used in measuring compliance with rules. However, rating scales are used widely in other types of program quality assessment systems—accreditation and research tools.

Weighting System

A Likert type of measurement tool that utilizes a modified Delphi technique to determine the relative risk to individuals if there are violations with specific rules. Weighting systems are developed by sending a survey to a selected sample of persons in order for them to rank the relative risk of violation with specific rules.

Outcome Based Systems

A measurement system based upon outcomes, not processes. A facility would be assessed by the outcomes it produced with individuals. For example, the number of consumers (children or adults) developing normally, free from abuse, not in placement, involved actively in the community, etc. are outcome based measures.

■ INSTRUMENT BASED PROGRAM MONITORING

Instrument Based Program Monitoring (IPM) is a particular approach to measurement and assessment. It is in contrast to a more qualitative type of assessment (case study is an example of this type of assessment). IPM is very quantitative and is characterized by the use of checklists (see the next section for a discussion of checklists). The advantages of instrument based program monitoring are the following: cost savings, improved program performance, improved regulatory climate, improved information for policy and financial decisions and ability to make state/province comparisons.

IPM is a paradigm shift in conducting licensing inspections and licensing of facilities.

It is an approach that lends itself to automation, it is objective and it is

generally systems-oriented. The IPM approach came into its own in the 1970s and has been used predominantly since then as the primary licensing measurement tool. Some individuals have argued that the IPM approach is not as effective as the more qualitative, narrative case study approach although they can't argue with its efficiency. A combination of IPM (quantitative approach) with a qualitative approach is probably most effective; however, this is very time consuming and a luxury that most state/province licensing agencies do not have, with more and more facilities to license and fewer and fewer staff to do the licensing.



■ CHECKLISTS

Checklists are the predominant means of collecting licensing data. It simplifies the process, making it very quantifiable. This is one of its strengths, but along with this simplification, a drawback is that some of the richness of the description of a particular facility is lost.

There are particular steps that need to be followed in the development of the checklist. Licensing administrators need to follow this four step process:

- 1) Make interpretations of the rules part of the overall manual for measurement of the comprehensive set of rules.
- 2) Identify the rules to be included in the checklist.
- 3) Consider the organization of the checklist—the flow of the investigation to the facility.
- 4) Decide what type of record keeping will be used—NCR paper, notebook computer in the field, etc.



■ RATING SCALES

Rating scales will not be discussed in detail because their applicability to licensing measurement is rather limited. Only in cases where a licensing administrator was interested in some form of partial compliance would rating scales make sense. The NAEYC (National Association for the Education of Young Children) accreditation system is one example of the use of a rating scale of full, partial or non-compliance with accreditation standards. While a partial compliance rating may be useful in accreditation standard measurement, it is generally not appropriate for use in licensing rule measurement.

Most licensing agencies do not use partial compliance, and the movement within the regulatory administration field is to consider partial compliance as being equivalent to non-compliance. Either a facility meets the rule or does not meet the rule. There is no middle ground.

■ WEIGHTING SYSTEMS

Weighting systems and licensing indicator systems that are described in the next section of this chapter are enhancements of the basic checklist (instrument based program monitoring) system. Weighting systems are used to increase the amount of variance in licensing compliance data. Because licensing data are nominal data ('yes' or 'no' compliance) and are generally highly in compliance, there is little variance in the data set from any particular set of rules. In order to increase the variance in data, weighting systems are used so that each rule does not have an equal weight. If you do not weight rules, by default, you have given an equal weight to each rule.

The remainder of this section describes the process for developing a licensing weighting system for use in the implementation of human care licensing rules, displays data from states that have used this approach and discusses the applicability of weighting systems for all types of human service licensing.

A licensing weighting system is a regulatory administration tool designed for use in implementing human care licensing rules. A licensing weighting system assigns a numerical score or weight to each individual licensing rule or section of a rule, based upon the relative health, safety and welfare risk to the consumers if a facility is not in compliance with the rule. The type of license issued is based on the sum of the numerical weights for each rule that is not in compliance.



The specific objectives of a licensing weighting system are:

- a) To standardize decision-making about the type of license to be issued
- b) To take into account the relative importance of each individual rule
- c) To ensure that rules are enforced consistently
- d) To improve the protection of consumers through more equitable and efficient application and enforcement of the licensing rules

■ WEIGHTING SYSTEMS

A licensing weighting system can and should be developed and implemented only if:

- 1) Regular or full licenses are issued with less than 100% compliance with rules. If a regular license is not issued unless all violations are corrected at the time of license issuance, a weighting system is not necessary. A weighting system is useful if a facility is issued a license with outstanding violations (and a plan to correct the non-compliance areas) at the time of license issuance.
- 2) There is a large number of licensing rules with a variation of degrees of risk associated with various rules. If there are only a few rules with equal or similar risk associated with each rule, a weighting system is not necessary. A weighting system is useful if there are many rules with varying degrees of risk.
- 3) A standardized measurement system or inspection instrument is used to measure compliance with licensing rules. Before developing a weighting system, a standardized measurement instrument or tool should be developed and implemented.

■ WEIGHTING SYSTEMS

Development of a Weighting System

This section will provide a step-by-step process in the development of a weighting system for licensing agency use.

- 1) The first step in developing a licensing weighting system is the development of a survey instrument. A licensing inspection instrument or measurement tool can be adapted into a survey tool. The survey should contain each rule or section of a rule, according to how it is measured in the inspection instrument. Survey instructions should explain the purpose of the survey and instructions for completing the survey instrument. It is suggested that survey participants rate each rule section from 1-8 based on risk to the health, safety and welfare of the clients if the rule is not met (1 = least risk; 8 = most risk).

The survey participant should be instructed to circle their rating choice of 1, 2, 3, 4, 5, 6, 7 or 8. An example of a survey question is:

Interior stairways, outside steps, porches and ramps shall have well-secured handrails.

1	2	3	4	5	6	7	8
Low Risk				High Risk			

- 2) Surveys should be disseminated to at least 100 individuals. If a state has more than 3,000 licensed facilities in the type of service being surveyed, consideration for surveying more than 100 individuals should be given.

Individuals surveyed should include providers of service; provider, consumer and advocacy associations; health, sanitation, fire safety, medical, nutrition and program area professionals; licensing agency staff including policy/administrative staff and inspectors; consumers of service; and funding agency staff. In order to assure a higher survey return rate, persons selected as survey participants should be contacted prior to the survey to explain the weighting system and request their willingness to complete a survey. *(See Karen Kroh's paper for detailed graphics of Pennsylvania's survey distribution.)*

■ WEIGHTING SYSTEMS

- 3) Survey results from each survey should be collected and entered into a computer data base spreadsheet software package. After all survey data are recorded, means or average weights for each rule or section of a rule should be calculated using SPSS—Statistical Package for the Social Sciences or SAS—Statistical Analysis System. (For detailed information on the statistical methodology employed in the development of weighting systems, see Griffin and Fiene’s *A systematic approach to child care regulatory review, policy Development of a Weighting System evaluation and planning to promote health and safety of children in child care: A manual for state and local child care and maternal and child health agency staff.*)

If there is sufficient variation in the means for each rule, the individual rule means can be rounded to the nearest whole number. Generally when comparing mean weights among the various groups surveyed there should be a similarity in rating among the groups, supporting the use of the weights as a reliable measure of risk.

- 4) The next step is to either (a) pilot test the weights with new licensing data for about six months or (b) apply the weights to at least 25% of historical data from the previous 12 months.

The intent of the pilot application is to collect data to use as the database for determining statistical cut-off points for the issuance of specific types of licenses or for administration of various negative sanctions.

A total weighted score for each facility based upon the combined weights of all violations should be calculated. Following is an example of how the scores should be calculated:

<u>RULE VIOLATIONS</u>	<u>WEIGHTS</u>
# 1	7
# 2	6
# 3	<u>+ 8</u>
Sum of Weights = 21	

■ WEIGHTING SYSTEMS

Under the above example a perfect compliance score with non-compliance areas would be a score of “0”. The higher the score, the lower the compliance would be. However, this is not congruent with the common usage of scores in which the higher score is associated with better compliance. In order to accommodate our familiarity with higher scores for the better facilities, the weighted score should be deducted from an arbitrary constant score of “100”. Thus a weighted non-compliance score of “20” will convert to a positive score of “80”. A facility with no violations will have a perfect score of “100”. This is more intuitive to individuals as they think about scores and measurement.

Using the previous example, the final weighted score would be computed as follows:

<u>RULE VIOLATIONS</u>	<u>WEIGHTS</u>
# 1	7
# 2	6
# 3	<u>+ 8</u>
Sum of Weights = 21	

Final calculation:

100
<u>-21</u>
79

■ WEIGHTING SYSTEMS

- 5) The fifth step in the process is to compute and apply the standard deviation or the median if the data are very skewed.

The mean and standard deviation of all final weighted scores computed in the pilot application in step #4 should then be calculated. Based upon experience with implementing licensing weighting systems, it is recommended that if a final weighted score is no more than one standard deviation below the mean, a regular license should be issued. If a score is between one standard deviation below the mean and two standard deviations below the mean, a provisional license should be issued (the length of the provisional license will vary based upon the severity of the non-compliance), or intermediate negative sanctions should be administered. If a score is less than two standard deviations below the mean, no license should be issued or a more severe negative sanction should be administered.

For example, if the standard deviation is 18 and the mean is 88, following is the distribution of the weighted scores used to determine the type of license to be issued:

Score of 100 — 70 = Regular license/no sanction

Score of 69 — 52 = Provisional license/intermediate sanction such as warnings, administrative fines or restriction on admissions

Score of 51 and below = No license/severe sanction such as revocation or administrative closure

- 6) The final weighted scores from the pilot application should be applied to the standard deviation cut-off points to determine the type of license or negative sanction issued. These data should be studied to compare types of licenses or sanctions issued under pre-weighting vs weighting.

■ WEIGHTING SYSTEMS

- 7) Before implementing the licensing weighting system the following additional licensing factors should be considered and incorporated as necessary into the licensing system.
 - a) repeated violations from the previous licensing inspection;
 - b) violation with high risk items (possibly a weight of 8.0 or above);
 - c) discretion of licensing inspector to recommend variance from licensing weighting system.
- 8) Whenever licensing rules are amended, or at least every 5 years, the weights should be recomputed and the weighting system re-evaluated.

The licensing weighting system as described here can be used to license any type of human care facility including child care, adult care, residential care and part-day care facilities. Licensing weighting systems have been developed in Pennsylvania, Utah, Florida and Georgia.

Since the concept, development and implementation of weighting systems is relatively new to the field of licensing, the long term impact and benefits of weighting systems have not been fully realized. The potential of using weighting systems and modifications of weighting, to help standardize the implementation and enforcement of licensing rules is an exciting area of research to pursue in the field of regulatory administration.

■ LICENSING INDICATOR SYSTEMS

As mentioned in the weighting system section of this chapter, indicator checklists or licensing indicator systems are used to improve upon instrument based program monitoring (checklist) systems. The licensing indicator system is one method of assuring compliance with licensing rules in a time efficient manner. The concept has been developed and successfully implemented in several states and for different human service types. The licensing indicator system was originally developed in Pennsylvania in 1977 for use in licensing child care centers. The original intent was to develop an abbreviated licensing instrument in order to refocus licensing investigation time to assess and assist in quality enhancement activities.

From 1980-1984, the US Department of Health and Human Services funded a project to study and further develop a licensing indicator system for child day care facilities on a national level. The federally funded project, known as the Children's Services Monitoring Transfer Consortium, organized researchers, state licensing administrators and professional staff from Pennsylvania, Michigan, West Virginia, Texas, New York City and California to review and refine the existing Pennsylvania system for possible use by other states.

The licensing indicator system is now used to assist in licensing human care facilities in Pennsylvania, West Virginia, Texas, Maryland, Utah, Florida, Delaware, Georgia, Washington, Minnesota and California.

The purpose of a licensing indicator system is to increase the efficiency and effectiveness of an existing licensing system by refocusing the emphasis of the licensing process. A licensing indicator system is intended to complement, and not replace, an existing licensing measurement system. Through use of the licensing indicator system, less time is spent conducting annual inspections of facilities with a history of high compliance with the licensing rules, and more time is spent a) providing technical assistance to help facilities comply with licensing rules and b) conducting additional inspections of facilities and agencies with low compliance with licensing rules.

■ LICENSING INDICATOR SYSTEMS

The licensing indicator system is actually a shortened version of a comprehensive licensing inspection instrument. A small number of rules are selected based upon a statistical methodology designed for this specific purpose. The licensing indicator system uses a measurement tool, designed to measure compliance with a small number of rules, that predicts high compliance with all the rules. If a facility is in complete compliance with all of the rules measured in the licensing indicator system, high compliance with all the rules is statistically predicted. It is critical to understand that the rules for the licensing indicator system are selected statistically (the statistical technique is called the phi-coefficient and generally is set at a p value of .01 or higher) and not based upon value judgement (arbitrary assignment, no basis from research literature), risk assessment or frequent rule violations. The rules are selected based upon an SPSSPC+ computer software package that compares violations of facilities with high compliance versus facilities with low compliance. The rules that are most often out of compliance in low compliance facilities and in compliance in high compliance facilities will be the indicator or predictor rules.

Prerequisites for implementing a licensing indicator system

Before developing and implementing a licensing indicator system it is important that the existing licensing system is comprehensive and well established. The following are prerequisites to implementation of an indicator system:

- 1) Licensing rules must be comprehensive, well written and measurable. Rules are the building blocks for any licensing system. If the rules are not well written and measurable a licensing indicator system should not be pursued. Also, if the total number of rules is small, a shortened inspection tool is not valuable.

■ LICENSING INDICATOR SYSTEMS

- 2) There must be a measurement tool designed to standardize the application and interpretation of the rules. A licensing inspection instrument designed to assure statewide consistency in the application of the rules is essential prior to implementing a licensing indicator system.
- 3) There should be a licensing weighting system designed to assess the relative risk to consumers if the rule is not met. This system may be a formal weighting system or a simple classification system which categorizes rules by degree of risk. An example of a high degree of risk to consumers would be the accessibility of heat sources or toxins. Having a signature in a record is an example of a low degree of risk to consumers.
- 4) At least one year of data on rule violations for individual facilities. These data are needed to enter into the computer software system in order to determine the rules that are the indicators or predictors of high compliance.

■ LICENSING INDICATOR SYSTEMS

How to develop a licensing indicator system

The basic steps to developing a licensing indicator system include:

- 1) Select facilities to be used in determining the indicators. If the total number of licensed facilities is less than 200, all 200 facilities can be used. If the total number of licensed facilities exceeds 200, sampling must be done. Generally, a sample of 100 facilities or 10% is acceptable. When selecting the sample, variables of size of facilities, geographic area, urban/rural, profit/non-profit, public/private and varied compliance levels or scores must be controlled.
- 2) Violation data for the sampled facilities is entered into a computer software system designed for this purpose (SPSSPC+ is recommended—consult with NARA consultant Dr. Richard Fiene for the necessary syntax and computer coding for doing the analyses).
- 3) A list of indicator or predictor rules, based on phi coefficients, that were the best indicators of high compliance will be calculated by the computer software system. These are the rules that are most often out of compliance in low compliance facilities and in compliance in high compliance facilities.

■ LICENSING INDICATOR SYSTEMS

Indicator system

Individual Rule	Groups	
	Overall High Compliance	Overall Low Compliance
In Compliance	X	
Out of Compliance		X

- 4) A small number of additional rules which are determined based on a licensing weighting system or relative risk are added to the statistically selected indicators. The purpose of this step is to assure face validity of the instrument. By adding a smaller number of carefully selected high-risk rules to the instrument, the licensing agency can be assured that critical rules are always measured.
- 5) In order to assure that full compliance with all the rules is maintained, five items selected at random should also be applied as part of the licensing indicator system. The final licensing indicator system instrument contains the indicator rules, high-risk rules and random rules. The total number of rules on an indicator checklist will vary, but will range from 20-45 items.
- 6) Specific criteria for use of the licensing indicator system are developed.

■ LICENSING INDICATOR SYSTEMS

Criteria for use of the licensing indicator system

The development of very specific criteria for use of the licensing indicator system is perhaps the most critical step of the design process. This is the step at which the determinations are made as to when the licensing indicator system will be used. The determination of use of the system should be standardized and not based upon licensing inspector discretion.

Each licensing agency must develop its own criteria based upon its own historical licensing data and experience. Following are some criteria that may be useful:

- 1) The facility has had a full or regular license and no negative sanctions have been administered, within the previous two (2) years.
- 2) The facility has had a score or percentage of compliance above a specified threshold for the previous year.
- 3) All previous violations have been corrected according to the facility's plan of correction.
- 4) No significant validated complaints have been found within the past year.
- 5) The total number of consumers served has not increased by more than a specified percentage during the past year.
- 6) There has not been significant staff turnover at the facility/agency within the past year. This may be targeted to certain levels of staff turnover, such as direct care staff or facility directors, depending on which staff are particularly key for program stability.
- 7) A full inspection using the comprehensive licensing measurement instrument must be done at least every three (3) years.

■ LICENSING INDICATOR SYSTEMS

Revision of the licensing indicator system

The licensing indicator system should be continually reevaluated for its effectiveness. The system should be completely revised at least every three years or upon a revision of the rules. In order to achieve the intended purpose of the licensing indicator system of refocusing the emphasis of licensing effort from facilities with high compliance to facilities with low compliance, constant review, evaluation and revision of the licensing indicator system is essential.

Other types of inferential inspection systems, of which the licensing indicator system is only one, will not be addressed in this chapter because inferential systems other than the licensing indicator system have not been determined to be statistically valid or reliable. As licensing administrators may potentially need to defend their actions in a court of law, it is essential that the methodology or technique utilized is scientifically sound. When it comes to inferential inspections only those instruments based upon an indicator or weighting methodology can stand up to this rigorous testing.

■ OUTCOME BASED SYSTEMS

This is a relatively new phenomenon in the licensing and regulatory administration field. The emphasis in this new approach is to examine outcomes rather than processes. What are the ultimate outcomes for individuals? Determine this and the argument goes—there is no need to measure processes directly.

Outcome measurement is appealing in many respects. It does focus on results, something the human services field was short on demonstrating in the 1990s. However, there is a fallacy in this approach. Results are the end product, but we always have a process to get to the end product.

Another issue is that the purpose of licensing is to prevent harm to consumers. A purely outcome-based system would potentially harm consumers who were in the facilities later determined to “fail” the outcome test. Moreover, there are two other problems:

1. Insufficient (political) agreement on what are acceptable outcomes.
2. Some outcomes will not manifest for years and/or are contaminated by other variables related to other influences on later behavior.

What makes more sense is to tie outcomes to specific regulatory processes that appear to be in a causal or at least a correlational relationship. If licensing agencies were able to clearly link specific results (outcomes) to specific rules (processes), there would be the empirical ability to focus only on those rules that produced positive results for consumers and families and eliminate all other unnecessary rules that do not produce positive outcomes for consumers and families. Specific studies could be conducted and in fact have already been conducted by university researchers. In child care, for example, low staff:child ratios, pre-service and in-service training of staff, highly qualified staff and small group size are all examples of regulatory variables that have been identified as surrogates to program quality that produce positive outcomes for children.

Outcome based or results-oriented systems will impact licensing, but the research literature demonstrates how licensing agencies can clearly link outcomes to regulatory processes that produce the outcomes. This becomes a powerful argument to legislators when this roadmap of process to outcome can be provided.

■ RELATIONSHIP BETWEEN RULES AND INSTRUMENTS

This section is included because this is one area that gets many licensing administrators into trouble. Not enough time is spent on making sure that the instruments developed are the exact reflection of the rules. This is where the interpretive rules that are part of any measurement instrument that accompanies the actual instrument should be placed. This helps to increase the reliability of the instrument and doesn't hurt the overall validity of the tool either (more on reliability and validity in the next section). Readers should refer to Chapter 2, *The Formulation of Rules*, for additional information on the definition and development of interpretive and substantive rules.

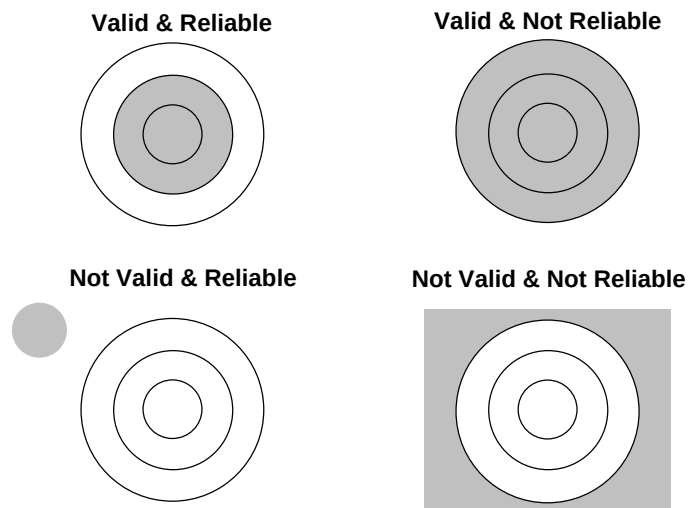
When there is not a close link between instrument development and rule formulation this only leads to headaches for licensing agencies. It may take years and not be evident until you get called into a court of law to defend your licensing system but it will happen.

The analogy of playing Russian Roulette may be useful. As licensing administrators, you are never 100% certain that all your facilities are compliant with all the rules. However, there are certain management procedures and processes that you can put in place to help. A clear link between rules and measurement tools is one of them. Since you are never 100% sure of full compliance (in other words all six chambers of the revolver are not empty—if they were, you wouldn't have Russian Roulette), you must make difficult decisions related to increasing or decreasing your chances in playing Russian Roulette. So you have the choice of having the management and procedural safeguards built in (one or two bullets in the revolver) or you don't build in the procedural safeguards (four or five bullets in the revolver). It is obvious statistically where your chances are greater in surviving a potential mishap in a licensing system.

■ RELIABILITY AND VALIDITY

The two concepts of **reliability** and **validity** are so critical to measurement, but are so often overlooked in the development of licensing measurement systems. In fact, it has been estimated that as many as 30 states may be using a type of inferential inspection. But only 1/3 of these states has followed the rigorous statistical methodology as outlined in the Licensing Indicator System section.

Validity and Reliability



Very simply, **validity** deals with content of the particular tool or instrument—does it serve the purpose for which it is to be used? Does it measure the rules accurately? Usually the answer to this question is easier for licensing administrators to answer. Since licensing measurement tools should be directly based upon rules, as explained in the previous section, there should not be much difficulty in establishing validity. When the tools are not based on the rules, that is when validity can be and should be called into question.

■ RELIABILITY AND VALIDITY

Reliability deals with the administration of the tool or instrument. Does it measure the rules consistently and in an objective manner? The answer to this question is much more difficult for licensing administrators to answer affirmatively. This poses real problems if each administration of the licensing tool is not consistent and objective. Facilities will not have the rules applied in an equal and fair manner.

Reliability testing should be done methodologically and scientifically. Inter-rater reliability should be established for the tools/checklists that are to be used in the field by licensing field staff. This is a process that has been well documented in the psychological research. This has not been the case within licensing and regulatory administration. Generally checklists are designed quickly and are never tested for reliability. This creates a problem that many of us have heard—the rules are not applied uniformly across the state/province. The reason is that the tool that is used to measure compliance is not reliable.

In order to establish reliability, licensing inspectors need to go out to facilities in pairs assessing the same facility at the same time. They then need to compare their results. Do they agree on what is in compliance and out of compliance at the particular facility? If there is not at least 90% agreement for each rule then additional interpretation of that specific rule is needed. Establishing reliability is not overly difficult nor overly time consuming; however, it will add a bit more time before staff are really ready to begin to license facilities (90% agreement on each rule and interpretative rule).

■ BALANCE BETWEEN COMPLIANCE AND PROGRAM QUALITY

An interesting development in the past five years has been the emphasis on program quality as a result of pressure from consumers, families, advocates and the general public. Consumers and other interested persons are requesting licensing agencies to ensure not only the health, safety and well-being of individuals served in facilities, but also to be concerned advocates for the overall quality of services provided at these facilities.

This increased emphasis and concern for program quality is a difficult area to address for licensing agencies. The resources to complete program quality reviews and to advocate for quality within government are not commensurate with the expectations. However, there are some strategies that can be employed to assist licensing agencies. The first and foremost will be to save time on doing licensing inspections. The indicator system described in this chapter will provide such a tool for saving time. Studies conducted over the past two decades indicate that utilizing an indicator checklist approach saves up to 50% in the on-site inspection time.



The time saved in doing licensing inspections should be used to either:

- a) Conduct additional licensing inspections in new or problem facilities
- b) Provide technical assistance
- c) Complete program quality reviews

This could be done by utilizing a tool from accreditation in observing classrooms, or utilizing a program quality tool from the research literature (for example, *Early Childhood Environment Rating Scale*). Licensing administrators need to be certain that they have a plan to utilize this extra time or the worst fears of licensing professionals could occur. Two potential scenarios could play out. One is that the time is used to do more and more licensing inspections utilizing the indicator system on more and more facilities. The worst scenario is that staffs are cut. If a state/province can complete all its inspections in half the time, then doesn't it follow that only half the staff is needed? With a clearly articulated plan on how the licensing and program quality reviews will produce higher quality programs should help to prevent this cost cutting approach. However, this is always a fear that licensing administrators must face.

■ CONCLUSION

This *NARA Licensing Curriculum* chapter provides a brief overview to the major issues confronting licensing administrators when they consider licensing tools and measurement systems. The emphasis upon quantitative systems was reflected in this chapter because of the need to develop cost effective and efficient licensing systems as the number of facilities continues to grow with shrinking resources. Also there is a compounding effect with higher expectations on licensing agencies to be concerned more about program quality.

The chapter showed the various types of measurement tools that apply to licensing and regulatory administration. It is clear that given the nature of licensing there are certain tools more suited than others, such as checklists versus rating scales. A very detailed description of both licensing weighting and indicator systems was provided. The reason for this emphasis is that these are two very valid and reliable tools that can be used by licensing administrators in making their agencies more effective and efficient. The licensing measurement field is changing constantly as new approaches are introduced. For example, within the program evaluation field there is a move to have a better balance between quantitative and qualitative analyses. It will not be long before this initiative has its impact on the licensing measurement field as well.

■ REFERENCES

Harms and Clifford. *Early Childhood Environment Rating Scale*. New York: Columbia University Teachers College Press, 1980.

Kroh, Karen E. "Development of a licensing weighting system for use in the implementation of human service licensing regulations," *Licensing of Services for Children and Adults: A Compendium of Papers*. Virginia Commonwealth University, October 1987.

Griffin and Fiene. "A systematic approach to child care regulatory review, policy development of a weighting system evaluation and planning to promote health and safety of children in child care: A manual for state and local child care and maternal and child health agency staff." *Zero to Three*. The National Center for Clinical Infant Programs, 1995.

The nagging issues of quality, accessibility, and affordability

Searching for a Solution to the Child Care Trilemma

by Richard Fiene

Every day we read about child care crises: Parents cannot find adequate care. There is not nearly enough quality child care. Qualified teachers are leaving for public school jobs where they can increase their salaries by 20-30%. Staff turnover is at 30-40%. Research tells us the majority of care in the United States is mediocre at best. All these issues point to the trilemma of quality, accessibility, and affordability that has been nagging at American child care for at least the past decade or two.

A solution to the trilemma equation in child care of quality, accessibility, and affordability has been difficult to address. In their campaign for adequate compensation for early childhood staff, the National Association for the Education of Young Children has documented the loss of the most highly qualified early childhood professionals to public school early childhood programs and to other professions. Because of low wages, early childhood staff cannot live on their teaching salary alone without supplementing it with other forms of employment. This is an impassioned issue because so much is at stake — staff-

child ratios, ability of parents to afford child care, and availability of sufficient care. Staff-child ratios, for certain, has been one of the sacrosanct surrogates of quality viz a viz the regulatory system and is the key to the solution of the trilemma. The research over the past 20 years clearly demonstrates the relationship between the number of children and the number of adults in a child care setting.

In the past as one alters the quality portion of the child care trilemma equation, this impacts both the accessibility and affordability portions of the equation. If the accessibility or affordability portions are altered in

any way, the quality portion of the equation is changed. There is a winner on one side of the equation but there are also always losers on the other side of the equation. There has not been a viable solution in which compensation can be increased to staff with no equivalent cost increase to parents, while at the same time increasing the number of children served. This article proposes a potential solution to this nagging problem.

A new concept (trilemma solution — tying compensation to staff quality without increasing cost to parents) is being proposed. This concept needs to be well researched, it is not one that state licensing administrators should think of in terms of making changes in policy at this point. There



Richard Fiene, Ph.D., a research psychologist specializing in early childhood program evaluation, is a research professor of psychology and education in the graduate school at the Pennsylvania State University and professor in charge of the graduate program in early childhood education at the Pennsylvania State University at Harrisburg. He has directed several national early childhood research and evaluation projects.

are too many issues related to waiving regulations, burnout of staff, and impact on children and teachers — short term and long term — that need to be ascertained before the policy implications are discussed. As a footnote to this proposal, this concept being proposed is for preschool care and not for infant or toddler care.

Trilemma Solution: A New Concept

The potential solution to the trilemma is to begin with the quality sector. Quality of the program is tied to staff quality and the number of staff to children (the staff line item is the most costly portion of a child day care budget as well). Higher education, direct training in early childhood or child development, and more years of experience generally correlate with a higher quality level of care. The more highly qualified staff a program has, the higher the quality of the overall program.

Most regulations address the adult-to-child ratio from an absolute (linear) standpoint. There is a specific ratio based upon the ages of children served. The adult-to-child ratio does not take into account any qualifications related to staff. In fact, most states cancel out the difference in education by equating it to experience so that the following scenario plays out:

Staff Qualifications			Children-to-Adult Ratio
Education	Experience		
AA	+	4 years	10-1
BA	+	2 years	10-1
MA		none	10-1

However, another spin on the above is the following example, a staff person with a master’s degree in early childhood, with 30 hours per year of in-service workshops and 10 years of experience cares for the same number of children as an entry level bachelor degree staff person, with 6 hours per year of in-service workshops and no experience:

Staff 1 = MA + 30 hours in-service +
10 years experience = 10-1 ratio

Staff 2 = BA + 6 hours in-service +
no experience = 10-1 ratio

If a state were to address the adult-to-child ratio from a relative (non-linear) standpoint, taking into account the qualifications of staff, a very different scenario could occur. For example, the following could occur (for ease of presentation, only educational qualifications and years of experience are addressed here):

Staff Qualifications	Year of Experience									
	1	2	3	4	5	6	7	8	9	10
AA		10-1			10-1->11-1					11-1
BA		11-1			11-1->12-1					12-1
MA		12-1			12-1->13-1					13-1

In the research literature, more advanced degrees by themselves do not necessarily correlate with a higher level of care. Direct in-service training in ECE/CD needs to be entered into the equation. (See Figure 1.)

The implications for such a model have tremendous cost and availability enhancements. On the availability side, as ratios go higher, more children can be served. As these ratios increase, more revenue can be brought into a program which can then be used to pay for the higher qualified staff person. By using this approach, however, no additional cost of service is passed on to the parents or the program. The unit cost stays the same, only more children per qualified staff person are served.

At a practical level, taking Figure 1 into consideration, how would this really work? Let’s take a classroom of 4 year olds — 10 children with a 10-1 ratio. The teacher has a master’s degree with 10 years of experience and has been taking continuing education credit. The teacher has an annual salary of \$20,000 per year. The unit cost for preschool care is \$3,500 per year. To implement the concept, the teacher with the master’s degree would be the individual we want to potentially impact in the following manner:

The ratio in the classroom would move from 10-1 to 11-1 with an additional 4 year old being allowed to enroll. It is assumed that there is sufficient space (35-40 square feet per child) for the additional child. It is also assumed that \$500 of the \$3,500 is for the additional cost related to having the child in the classroom. The remaining \$3,000 would go to the teacher as a permanent salary increment (the center would have to agree to this) — the teacher’s salary would go from \$20,000 to \$23,000 per year. This would be a 15% increase in salary.

By using the relative adult-child ratio as stated above, taking quality of staff into account when determining ratios, this model could provide a potential solution to the child day care trilemma of quality, accessibility, and affordability. Quality increases by having more qualified staff in those classrooms with lower ratios.

It could be argued that by having lower ratios, quality will be lowered as well. This has been demonstrated in the research literature. However, with the model presented here, this would only occur when the most highly qualified staff were in these classrooms. Higher ratios would have to be maintained in those classrooms with

Figure 1
Qualifications and Training Tied to Compensation

Staff Qualifications	Years of Experience or Number of Training Courses									
	1	2	3	4	5	6	7	8	9	10
	Resultant Ratios									
AA	10-1				10-1->11-1 + \$3,000 salary increase					11-1
BA	11-1 + \$3,000				11-1->12-1 + \$3,000					12-1 + \$3,000
MA	12-1 + \$3,000				12-1->13-1 + \$3,000					13-1 + \$3,000

No additional cost would be charged to parents. Compensation for staff increases are totally from the additional children served per classroom.

staff who have lower qualifications. More children in the end could be served. Program income would increase. The additional dollars would go to pay the higher qualified teacher. This would also help to promote a professional development system. The more highly trained, experienced, and educated teachers would be paid a higher salary based upon the additional children. Parents, however, would not have to pay more because the additional income is from more children rather than a higher unit cost.

Conclusion

As dollars become tighter, more creative regulatory policy based upon research will need to be employed. This model takes into account the latest early childhood research and suggests a revision in how states' regulatory policies related to staff-to-child ratios are determined. Research clearly shows the linkage between the quality of programs being directly influenced by staff quality and number of staff to children. This model takes this into account and addresses several issues related to affordability and accessibility at the same time. (See Figure 2.)

This concept is one that needs to be fully researched. Hopefully, researchers, center based administrators, and state policy administrators can partnership together. This concept has as many questions as it does promise and potentially as many drawbacks if not well researched. As stated earlier, and I want to emphasize this, this is not a suggestion for state licensing administrators to begin to waive staff-child ratio regulations and make this state policy. It is suggested, however, that on a limited basis within a research context this concept be tested to examine the benefits and the drawbacks. Will this impact staff turnover?

Will the additional dollars be sufficient to keep our most qualified early childhood teachers in child care?

As a final footnote or afterthought to what has been proposed in this article, I want to be very clear that this proposal is an intermediate solution but not a long-term solution to solving the trilemma in child care. This is a very controversial proposal. I have had professionals argue passionately on both sides. However, given the present state of economics, I see this as a solution to hopefully keep our most qualified staff in child care until additional dollars can be found. Increased compensation not tied to staff-child ratios is the solution, but I do not see that happening realistically in the near future.

Figure 2
Child Day Care Trilemma's Potential Solution

- ✓ Links training to compensation
- ✓ Develops a professional development system
- ✓ No additional cost to parents
- ✓ Links training to quality
- ✓ Ties quality to regulations through increased responsibility
- ✓ Links quality to accessibility and affordability



Pergamon

Journal of Early Childhood Teacher Education 22 (2001) 267–274

*Journal of Early
Childhood
Teacher
Education*

Mastering course content and learner satisfaction in early childhood education: A comparison of regular classroom instruction with three variations of internet delivery☆

James Johnson^{a,*}, Richard Fiene^a, Jane Keat^a, Harriet Darling^a,
Donald Pratt^b, Joyce Iutcovich^b

^a*The Pennsylvania State University, Pennsylvania, USA*

^b*Keystone University Research Corporation, USA*

Abstract

In the spring semester 2000, a Penn State course, ECE 479 (The Young Child's Play as Educative Process), was taught by the same instructor in four delivery formats. One group consisted of a regular classroom, held on campus. A second group, also on campus, was taught in a computer lab via the Internet; and there were opportunities for interaction with peers and the instructor. A third group took the course on the Internet as part of a local distance education group; hence, there were some limited opportunities for face-to-face interaction with peers and the instructor. The fourth group took the course on the Internet, as part of a statewide distance education group, where there were no opportunities for face-to-face interaction. Twenty students who enrolled in the course (5 per group) completed questionnaires and phone interviews. Information was gathered on professional backgrounds, computer experience, and initial level of content knowledge on the topic of the ECE Internet course. Sixteen students who completed the course were interviewed again to evaluate satisfaction with the course and to estimate learning outcomes. Across the four conditions general satisfaction was expressed with the content, activities, and course requirements and with the teacher. However, students in the three computer groups expressed dissatisfaction over technical problems (all four who did not complete the course came from these computer groups). Significant gains in content knowledge occurred for the classroom group, while the learning in the three Internet-based instruction groups did not show the same gains. Concern was expressed related to the lack of face-to-face interaction, making the learning environment less desirable. Although Internet technology provides a great deal of promise, these results suggest that improvements are needed to make this delivery modality more effective for in-service distance learning. © 2001 Elsevier Science Inc. All rights reserved.

1. Introduction

"I think the thing that I liked least and I know it's inherent in an Internet course is lack of human involvement. Lack of human contact. Not getting to

talk with the instructor at all, face to face. That was very tough for me. The other thing was the lack of the immediate feedback. If you had a question you had to type it in and wait for the instructor to log on and read it and then get back to you."

☆This research was funded by Keystone University Research Corporation and the Penn State Consortium for Children, Youth, and Families. The authors gratefully acknowledge the role played by ECE doctoral students Hey-Jun Ahn and Juli Graboski in helping develop the on-line course and

World Campus specialist Cathy Holsing for her technical support and encouragement.

* Corresponding author. Tel.: +1-802-814-865 ext.2230; fax: +1-814-863-7602

E-mail address: jej4@psu.edu (J. Johnson).

"My frustration was that I don't like to fall behind in anything. I felt like I was falling behind because I wasn't able to access these readings, wasn't able to do some of the things on the computer that I should have been able to do. I was just getting frustrated and feeling like this isn't worth it for me. And that's why I dropped the course."

"The lack of being able to talk to anybody. The fact that you were doing all of this reading on your own. That was hard. I was used to sitting in classes and having someone lecture to me and then having students to interact with. It was difficult to get used to just reading it and absorbing it by yourself."

"Just simple things like just being able to print the page on my screen out. I didn't, I had no clue how to transfer stuff onto the hard drive and to be able to print something. I'm used to pressing the print button and stuff popping out. I didn't know it was that much involved."

The Internet is a new and ever expanding tool for learning (Benson & Meyers, 2000; Lan, 1999; Owston, 1998), including distance education (Williams, 1996). Whether or not the Internet can be an *effective* tool for training staff within the human services, particularly childcare staff, needs to be explored. Our project evaluated the effectiveness of the Internet training in terms of learning outcomes, its implementation (specifically, the technological aspects), and the student's level of satisfaction with the course. The impetus for our project came from two sources.

First, Pennsylvania delivers childcare training to all licensed and registered childcare providers in the state and is interested in making this system more cost effective and efficient. Discussions related to utilization of the latest technologies, such as the Internet, are being considered. Secondly, a new initiative, CyberStart, will link all licensed childcare centers in Pennsylvania to the Internet. While this initiative is specifically designed to offer Internet access and educational programming for children, it will also make this technology available to childcare staff. Hence, there is a need to evaluate the feasibility and effectiveness of Internet-based distance education as it becomes more available to childcare centers.

2. Research design

A quasi-experimental design was employed, which consisted of four groups of five students who enrolled in the early childhood education (ECE) course ECE 479 *The Young Child's Play as Educative Process*, which focused on play, communication, and curriculum. The first group experienced the traditional lecture/discussion course format. A second

group took the course on the Internet, but within the context of a computer lab (located at Penn State York) where they had the opportunity to interact with their peers and the instructor. A third group took the course on the Internet as part of a local distance education (DE) group; this group also had some face-to-face interaction with their peers and knew the instructor. A fourth group took the course on the Internet, but as part of a statewide DE group. This group had no face-to-face interaction with their peers or the instructor.

This research design enabled us to examine the available technology to determine any hardware or software constraints, as well as the efficiency of the technological support services, by comparing groups that took the course via the Internet in different environmental circumstances (i.e., the on-campus computer lab vs. a home-computer set up). The research design also permitted us to evaluate the importance of the human element as a component of the effectiveness of this training modality since participants were in controlled settings with varying possibilities for face-to-face interaction.

Evaluating whether Internet training delivered to childcare staff is an effective training modality has special significance at this time given the paucity of empirical evidence to support or refute such claims. Research is needed to help the Pennsylvania Childcare and Early Childhood Development (PA CC/ECD) Training System, and comparable systems in other states, make more informed decisions about the use of Internet-based distance education for teachers of young children.

Given the exploratory nature of this study, a qualitative approach with a small sample was employed to generate data from questionnaires and interviews, as well as from course assignments completed by students. The questionnaire included items that tapped demographic characteristics (age, sex, and prior education), current position and experience within childcare, and experience with computers. Phone interviews, administered before and after the course, and lasting in duration from 30 to 45 min each, assessed students' knowledge about play and perceptions about the course. In addition, select course assignments were independently graded to assess knowledge.

The course aimed to increase the students' knowledge about play and its practical application. Two measures of the learning outcomes were used: (1) interview responses to questions about play given before and after the course and (2) grades on selected course assignments. The latter were independently evaluated by faculty other than the course instructor. Both measures were scored without knowledge of group membership.

From the phone interviews, students' answers to four questions about play were evaluated. The questions were: (1) What is play? (2) What is the value of play? (3) What is positive play? and (4) How can adults have a positive influence on play? Content analysis revealed that numerous ideas were elicited by these four questions both before and after the course. An empirically based coding system was developed and employed to score students' answers. Appendix A provides the categories and subcategories used in scoring.

3. Course development

The course *The Young Child's Play As Educative Process* (ECE 479), offered by the Department of Curriculum and Instruction in the College of Education at The Pennsylvania State University, was selected for this experimental study of Internet-based instruction. This particular course was chosen because it was viewed as well-developed and because it included both theoretical as well as practical content relating to ECE curriculum and instruction. Developing the course for Internet-based instruction required connecting with the Penn State World Campus (online learning). A contractual relationship was formed involving Keystone University Research Corporation (KURC), the contractor for the PA CC/ECD Training System, the College of Education, and the World Campus. Technicians from the World Campus, in collaboration with ECE faculty and doctoral students, worked to develop the course in time for offering it in the Spring of 2000.

3.1. Stage one – course structure

One of the first tasks was to establish the course syllabus, which included the course objectives and requirements, along with a schedule and sequence of learning activities. Objectives included acquiring a knowledge base of theory, research and practice concerning play and ECE, improving play observation and documentation skills, analyzing playthings for use in ECE settings, planning suitable indoor and outdoor environments for play, understanding and making sound judgments about the use of play facilitation strategies, and learning how and why to become an advocate for play in education. Requirements included an observation project, designing play environments, writing letters explaining play-based teaching to parents and to a 'blue ribbon committee' of educational professionals, as well as doing an implementation project and keeping a journal. *The Instructor's Manual to Accompany Johnson/Christie/Yawkey Play and Early Childhood Development*,

Second Edition (Johnson, 1999) provided the guidelines for the overall course organization and sequencing, which followed the chapters of the text, with the content going from theory and research to policy and practice.

3.2. Stage two – course content

The course was organized into four modules with a number of online lessons or sessions in each module. There were a total of 26 sessions. Sessions were designed to last 50 min; students had a reading assignment for each session and a self-administered objective-item exam, which produced computer-generated feedback for self-evaluation. Also, open-ended discussion questions were assigned for chat room or bulletin board discussion. The objective items and the open-ended discussion questions came from the instructor's manual. Activities and problems were embedded in sessions and were sequenced and used in the same way across the four groups of students.

3.3. Stage three – course programming

The final formatting and programming of the course for Internet instruction occurred during the Fall of 1999. The course program included some special features to make the on-line learning experience more interesting. For example, an animated pop-in character (a cartoon owl) appeared on screen at various selected points throughout the sessions to ask questions as a real classmate might. The World Campus technical staff, with the assistance of the course instructor (third author), also prepared a home page for the course, which included the course schedule and contact information.

4. Participants

All 20 students were female, ranging in age from 23 to 60 with mean of 39 years. Sixteen students had a bachelor's degree and one a master's and two had a high school diploma or GED certificate. The most common college major was elementary education ($N = 7$). Only three students majored in ECE and had taken any previous courses on play. For the entire group, inservice training hours on this subject ranged from zero ($N = 5$) to 45 hours ($N = 1$), with a mean of 10.5 hours.

Fifteen students held the position of center director and three were assistant directors. One student had the title of childcare coordinator, and another was a personnel training coordinator and an assistant group supervisor. The length of time that the students worked at their current childcare center ranged from

a minimum of six months to a maximum of 23 years with the mean 9 years. The length of time that students had worked in the early childhood field ranged from 2 years to 25 years, with mean 11.9 years.

5. Findings and discussion

5.1. Prior computer experience

Students reported prior experience using personal computers ranged from no experience to 15 years. The mean value was 4.1 years. Seventeen students had a personal computer at home and were asked to indicate which specific activities they had used. Those uses included: word processing (15 responses); recreational software (12 responses); Internet (12 responses); and spreadsheets and/or databases (4 responses).

A questionnaire item asked who usually provided technical support for their home PC (e.g., installing new software or hardware, answering software questions, and fixing problems). No more than five students usually relied on themselves for technical support for their home PCs. Of the four students who chose "other," three students usually relied on a friend, and one usually relied on another teacher in her childcare center.

5.2. Knowledge about play

For the following four play questions, each student's answers given in the postcourse interview were compared with the precourse answers in order to evaluate whether there were response improvements, defined by an increase in positive responses and/or a decrease in negative responses (items 1 and 3), by an increase in the total number of distinct acceptable benefits of play given (item 2), or of distinct teacher's roles (item 4) cited.

5.2.1. Question 1: what is the definition of play?

Responses were coded into the categories and subcategories and summed into two groups. *Positive responses* tallied the number of Attributes, Distinctions, and Realizations (A, D, R); while a second group of *negative responses* summed the total of overly-inclusive, mistakes, slogans, and vague replies (X, M, S, V). Other codes were considered "neutral" answers and were not used in this evaluation.

All classroom and computer lab students improved, as did the three statewide DE students who stayed in the course. Of the three remaining local DE group students, one failed to exhibit a positive re-

Table 1

Pre- and post-course mean scores for benefits of play across four conditions

Groups	Classroom	Computer lab	Local DE	Statewide DE
Pre-course	5.2	4.0	3.8	3.0
Post-course	8.4	7.2	3.3	5.6

sponse at both times of measurement; they gave five negative responses at precourse time of measurement and three negative responses at postcourse time of measurement.

Classroom students gave a total of 25 positive responses after the course compared to only 3 positive responses before the course. This averages five positive responses per student and exceeds the averages on this index for the other three groups (1.2, 1.6, and 2.3 for lab, local DE, and statewide DE, respectively). Classroom and lab students showed few negative responses before or after the course, while local DE students gave 8 negative responses and statewide DE students gave 13 negative responses before the course. However, in postcourse interviews the six DE students made only five negative replies.

5.2.2. Question 2: what are the benefits of play?

Total scores for number of distinct benefits were derived from the coding system employed in the study. All general and expressive answers given (G and E) were added with all functional answers to yield a total benefits (G&E + F) score. Benefits of play results are shown in Table 1.

As can be seen in Table 1, the classroom and lab students performed better than the local DE and statewide DE students on this item. They were able to identify significantly more functions of play, especially after the course was taken. Nine of ten students in the classroom/lab groups improved, but only 3 of 6 in the DE groups did. Note also slight decrease in scores by local DE group.

5.2.3. Question 3: what is good play?

Here *positive responses* were the number of Nominal and Process (N and P) ideas from each student, while a second group of *negative responses* summed the total of Failure (F) responses exhibited. Other codes were considered "neutral" and were not used here.

Only seven of 16 students improved. Evidently, the course did not impact very much how well students could answer this question. The best answer, the P answer, was given by only three students, all at postcourse time. Failure to distinguish good play from play (e.g., "all play is good") happened six

Table 2

Mean scores for total roles across groups before and after the course

Groups	Classroom	Computer lab	Local DE	Statewide DE
Pre-course	3.8	2.4	4.0	3.6
Post-course	7.0	4.6	3.3	5.6

times at precourse time and eight times by the end of the course. Most students talked about types, functions and characteristics of play and failed to differentiate good play from play in an acceptable manner as determined by the coding and scoring system used in this study. This discouraging result may be because the scoring criteria were too harsh, or perhaps because interviewees did not understand the question. Also, it is possible that the course as taught, as well as the textbook as read, did not highlight the distinction well enough. In previous courses, this question has proven to be a "power question," distinguishing the "gold from the dross" in student performances.

5.2.4. Question 4: how can adults make child's play better?

Responses were analyzed with scores derived from the coding employed in the study. All general or attitudinal answers (G) per student were counted. Another group of answers covered the sum of all adult role responses (R). Total roles (G&R) were examined in the analysis. Table 2 shows mean scores across groups.

There were 11 students who improved, with all but one of them coming from the classroom or computer lab research groups. Only one student in a DE improved. Five students in DE groups actually performed more poorly in answer to this question after the course was over. In sum, the course influenced students' performance in identifying adult roles in children's play differentially depending on which research group they were in—classroom and lab students outperformed DE students.

5.3. Assessment of course requirements

There were three assignments that were graded for all the students: implementation activity (A), a parent letter (B), and a blue ribbon letter (C). As shown in Table 3, a comparison of the grades on these assignments across all four groups did not reveal that any one group consistently scored higher than the others. However, the traditional classroom group did score the highest on assignment B and had

Table 3

Scores on course assignments across four groups

Groups	Assignment A	Assignment B	Assignment C	Total
Classroom	2.86	3.86	2.75	9.47
Lab	1.90	2.63	3.00	7.53
Local DE	2.69	3.37	2.50	8.56
Statewide DE	2.99	3.61	2.35	8.95

the second highest set of scores on assignments A and C.

For assignments A and B the differences in scores for the four groups were statistically significant (Assignment A, $F=5.574$, $p < .012$ and Assignment B, $F=4.628$, $p < .023$). No significant differences were found for Assignment C. The computer lab group scored the lowest on assignments A and B while the other three groups of students all scored about the same. When all the assignments are totaled for an overall score, the traditional classroom performed significantly better than the other three groups ($F=5.221$, $p < .015$).

Students were also asked a series of questions about their expectations regarding the course, the benefits from taking the course, and their experiences in taking the course. The majority of the responses were very positive about the course in general, but the students who had taken the course via the Internet experienced a number of computer problems in either getting online or with the software. Everyone experienced problems and this delayed the start of the course for them for several weeks. There were problems with signing on, with the chat room, with passwords not being accepted, and so forth. However, once these problems were worked out, the course started and proceeded fairly well.

In sum, respondents revealed general satisfaction with the content of the course, the course activities, and course requirements. Moreover, there was a strong appreciation and high evaluation for the teacher. On the other hand, almost everyone (one or two exceptions) in the three computer groups expressed dismay over serious and continuing technical difficulties throughout the entire semester. However, even with a good course content, instructional design, teacher, and technical delivery, there were several people who clearly indicated reservations about Internet learning because it lacked face-to-face interaction. These people indicated that even if there were no technical difficulties, they would miss the human contact and would prefer courses or training taken in a classroom where there was greater opportunity for interpersonal interaction and contact.

6. Conclusion

This evaluation provides insights into offering ECE courses over the Internet. Clearly, it seems that the success of this technology is dependent upon the persistence and knowledge of the student for learning to occur. Four students who did not complete the course were from the local ($N = 2$) and statewide ($N = 2$) Internet-based distance education groups; none were from the traditional classroom or computer lab settings. The students were administrators of childcare programs, fairly familiar with computers and knowledgeable of the course content. However, even with these pluses, the students still had considerable difficulty in accessing and doing the course online. Possibly the dropout rate would have been greater if the students were at a beginning stage of their career (Cohen, 2000; but see Schrum, 1992). Students in the traditional classroom and in the computer lab groups, where there was more face-to-face interaction, scored the best on the interview play evaluations. Evidently, these two training modalities were more effective than were the local and statewide distance education training modalities where there was little or no face-to-face contact with other students or with the instructor.

On the three course requirements, the classroom group scored the highest on assignments summed together (9.47); but the computer lab group scored the lowest on the assignments (7.53), with the two distance education groups in the middle. This is inconsistent with the results from the analyses of the interview responses to the four questions about play. Here both the classroom group and the computer lab group gained the most. Although it is perhaps encouraging to see that the students in the two distance education groups scored higher on the course assignments compared to the computer lab group, their results were still lower than the students' scores from the traditional classroom group.

What have we learned from this study? For Internet instruction to be effective it seems that students must be technologically literate and knowledgeable about course content to some degree, and they must be persistent and highly motivated. Without these personal characteristics, the Internet course experience may not be a positive learning experience. The results of this study further suggest not utilizing Internet training across-the-board for childcare staff. It needs to be very targeted, beginning with directors of programs, who generally have the greatest experience

and education and potential exposure to computer technology. More generic, across-the-board training for the beginning level practitioner does not make training sense at this point.

Internet technology provides a great deal of promise for reaching childcare staff with needed specialized in-service training in ECE. But first it is necessary that the technology gets the fine-tuning to ensure its effectiveness as a training modality for the majority of childcare providers. Also, what is best for on-line learning? Professional attention must be given to criteria for deciding the kinds of training content (e.g. learning facts like state regulations or understanding and applying concepts like play in ECE) in relation to its packaging (e.g., a full course like in this study, or modules, or single sessions analogous to workshops). Setting up realistic expectations for learners and designing the right means of instruction and assessment are both very important. Finally, we recommend that priority be given to designing orientation modules to prepare potential users for Internet-based distance education.

References

- Benson, M., & Meyer, J. (2000). Student teachers' reflections and collaborations through electronic mail discussions. *Journal of Early Childhood Teacher Education*, 21, 123-127.
- Cohen, L. (2000, April). *Are we ready to abandon the classroom? The dark side of web instruction*. Paper presented at the annual meetings of the American Educational Research Association, New Orleans.
- Johnson, J. (1999). *Instructor's Manual to accompany Johnson, Christie, and Yawkey Play and Early Childhood Development* (2nd edition). New York: Addison, Wesley, Longman
- Lan, J. (1999, March). *The impact of Internet-Based Instruction on Teacher Education: The "Paradigm Shift."* Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, Washington, DC.
- Owston, R. (1998). *Making the link: Teacher professional development on the internet*. Portsmouth, NH: Heinemann.
- Schrum, L. (1992). Professional development in an informational age: An on-line experience. *Educational Technology*, 32 (12), 49-53.
- William, A. (1996, June). *Integrating courses with the internet: Preparing the teacher as well as the learner*. Paper presented at the summer conference meetings of the Association of Small Computer Users in Education, North Myrtle Beach, SC.

Appendix Table 1

Question 1: What is play?

- A1 Attribute of play given. For example, play is process-oriented, marked by positive affect, often non-literal, active, intrinsically motivated, etc.
- A2 Tangential attribute of play given. For example, play is something adults like to see in children; child and adult do play of own free will, etc.
- E Example of play is given. For instance, play is like playing with dolls or blocks.
- F Function of play is noted as a way to define it. A way to learn things. A way to release feelings.
- C Context of play is noted. Play can be solitary or done in a social group.
- X Overly inclusive or general statement about play is made, such as "it is a creative process."
- D Distinction is made, contrasting play with other similar behaviors such as imitation or exploration.
- R Realization is indicated, such as play is multidimensional, complex, or hard to define.
- T1 Type of specific play is given, such as constructive or dramatic play.
- T2 Type of general play is noted, such as educational play, dark play, and recreational play.
- M Mistake is made, a falsehood is uttered.
- S Slogan is cited; such as play is the child's work.
- V Vague pronouncement is made, unclear or hard to decipher.

Question 2: What is the value of play? What are its benefits?

General (educational) values include:

- G1 Teaches skills and abilities or is a learning experience
- G2 School preparation
- G3 Allows for the practice of skills
- G4 Generates further learning and development; allows children to go as far as they can with what they are learning
- G5 Avenue for creativity
- G6 Creates a well-rounded child
- G7 Allows child to explore and discover on one's own
- G8 Allows child to experiment

Expressive values include:

- E1 Enjoyment, fun, makes happy, love of life
- E2 Relaxation, release energy, tension
- E3 Vent frustration
- E4 Be self
- E5 Express self

Functional values include:

- F1 Cognitive: abstract thinking * imagination * learning content * creativity * learn on own terms-relevancy, meaningfulness * problem-solving * meta-cognition * memory * social cognition, empathy, perspective-taking * theory of mind * sense of self * sense of others * assimilation, integration, application of learning
- F2 Affective: motivating, feel good about self * self empowerment, sense of control * reduce anxiety, therapeutic, cathartic * self confidence, sense of self-assuredness
- F3 Social: learn to resolve conflicts * cooperation * group cooperation, team member * leadership skills * learn to share * learn to take turns * learn to help
- F4 Physical: gross motor * fine motor * learn to challenge self physically * self-help skills
- F5 Attentional: concentration * attention regulation * persistence
- F6 Assimilation
- F7 Language: communication skills * literacy * become good story-teller * vocabulary
- F8 Academic: reading and writing * shapes for math * science
- F9 Life skills, careers

Question 3: What is good play? Positive play?

- N Nominal answer is given, like educational play is good play.
- P Process of play is said to be important, "perking along", playing up to capacity.
- C Characteristic of play in general is given without really answering the question.
- T Type of play in general is given without really answering the question.
- F1 Function of play in general is given without really answering the question.
- F2 Failure to provide sensible reply, such as, "all play is positive."

(continued on next page)

Appendix Table 1 (continued)

Question 4: How can adults make play better for children?

There are many roles to perform in the ECE profession with respect to children's play. Some are general or attitudinal such as:

- G1 Value play, allow it to happen, be aware of your biases and those of others
- G2 Realize when attitude change about play occurs in self or others
- G3 Show an interest in play as a matter of public or educational policy
- G4 Lobby and advocate on behalf of children's right to play as an integral part of childhood

Other roles in the ECE profession with respect to children's play pertain to classroom or home or child development center situations where young children learn and are cared for. These include:

- R1 Stage manager, set up over-all positive and attractive environments, provide materials, playthings, space, toys, storage areas, time for play, literacy props, allow children choice of areas and activities, organize settings, rotate toys
- R2 Be a careful and systematic observer, evaluate play, change own behavior as a result of observations, document play process and products
- R3 Provide preparatory experience, bring in guest visitors, field trips
- R4 Play facilitator, scaffold, support, challenge children at play, do not be overbearing, ask questions, offer comments, suggestions, ideas for play, teach play
- R5 Co-player, be a play leader, model play behavior
- R6 Supervise play, make sure play is safe, monitor activities, mediate conflicts, enforce rules, help children negotiate, guide and discipline, be a referee

Other roles inherent in the ECE profession connected with play have to do not with the children but with other adults. These include:

- R7 Help adults be comfortable with play of children in educational settings, train teachers, educate staff about play in ECE
- R8 Inform and involve parents in quality play with young children in formal and informal educational settings

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

Richard Fiene

*Office of Children Youth and Families
Commonwealth of Pennsylvania*

Mark Nixon

Children's Services Monitoring Transfer Consortium, Washington, D.C.

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

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

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

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

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

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

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

Forces Changing the Regulatory Environment

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

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

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

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

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

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

IPM as a Response to These Forces

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

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

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

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

PROGRAM QUALITY SUB SCALE

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

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

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

TABLE 1
Pennsylvania Child Development Program Evaluation
Specific Items Within Identified General Areas

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

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

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

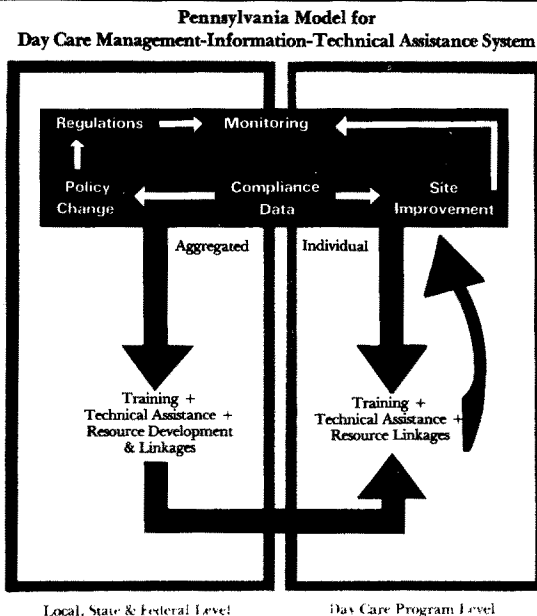


FIGURE 1

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

Indicator Checklist Improves IPM Systems

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

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

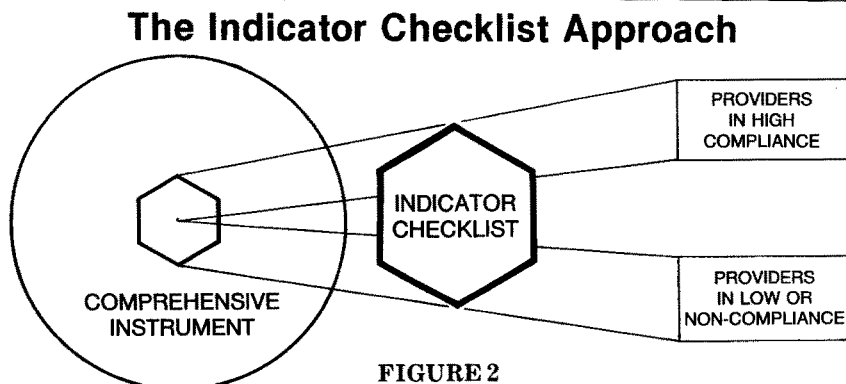


FIGURE 2

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

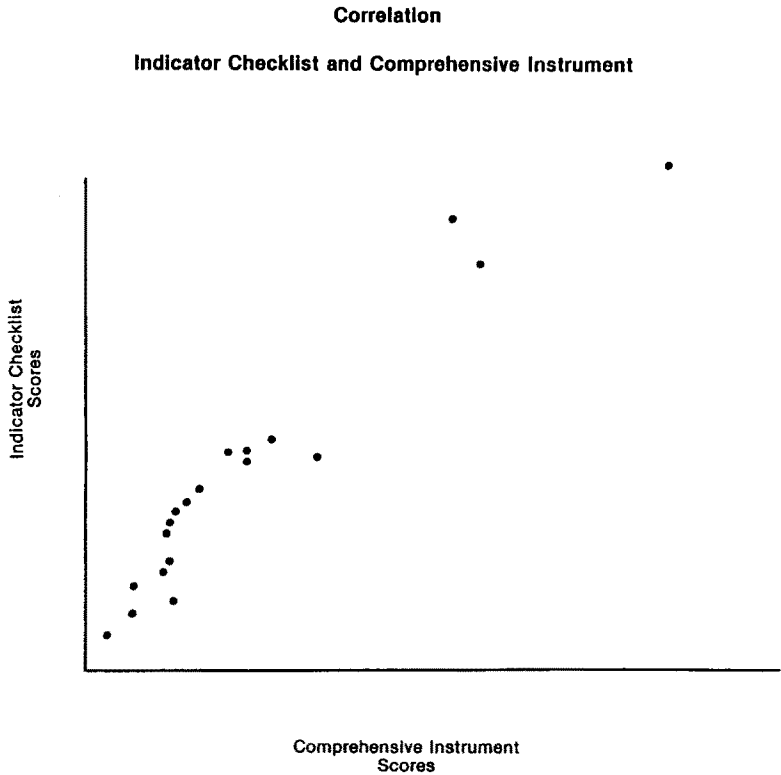


FIGURE 2a

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

LICENSING SUBSCALE

- A. GROUP SIZE AND ADULT CHILD RATIOS;
 - INFANTS 1 STAFF TO 5 CHILDREN
 10 INFANTS IN A GROUP
 - TODDLERS 1 STAFF TO 4 CHILDREN
 8 TODDLERS IN A GROUP
 - PRESCHOOLERS 1 STAFF TO 10 CHILDREN
 20 PRESCHOOLERS IN A
 GROUP
 - SCHOOL AGE 1 STAFF TO 15 CHILDREN
 30 SCHOOL AGE CHIL-
 DREN IN A GROUP
- B. SUFFICIENT SPACE—MINIMUM OF 40 SQ FT PER CHILD;
- C. EQUIPMENT IS EASILY ACCESSIBLE TO CHILDREN;
- D. ALL VEHICLES ARE EQUIPPED WITH AGE-APPROPRIATE SAFETY CARRIERS;
- E. CLEANING MATERIALS ARE INACCESSIBLE TO CHILDREN;
- F. EMERGENCY CONTACT INFORMATION IS AVAILABLE FOR ALL CHILDREN;
- G. ALL STAFF HAVE HAD PERIODIC HEALTH APPRAISALS;
- H. ACTIVITIES PROMOTE: DEVELOPMENT OF
 SKILLS
 SELF-ESTEEM
 POSITIVE SELF-IDENTITY
 CHOICE OF ACTIVITIES.
 (Fiene, 1984)

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

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

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

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

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

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

What is Needed to Develop an Indicator Checklist?

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

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

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

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

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

Specific Concerns of Administrators and Policymakers

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

Constructing The Indicator Checklist

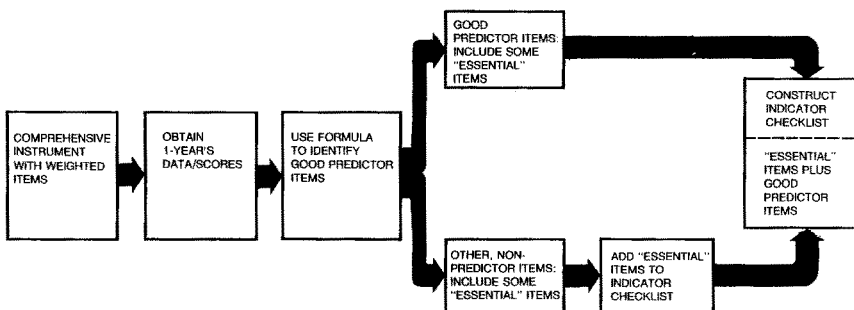


FIGURE 3

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

Can My State Make Use Of An Indicator Checklist?

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

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

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

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

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

Have Indicator Checklists Been Tried By Other States?

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

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

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

How Can The Quality Of Monitoring Be Ensured?

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

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

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

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

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

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

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

What Are The Potential Drawbacks?

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

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

TABLE 2

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

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

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

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

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

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

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

Conclusion

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

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

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

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

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

References

- Aronson, S., Fiene, R. & Douglas, E. *Pennsylvania Child Development Program Evaluation Instrument*, Bureau of Child Development, Harrisburg, Pennsylvania, 1977.
- Belsky, J. & Steinberg, L. The effects of day care: A critical review, *Child Development*, 1978, 49, 929-949.
- Class, N. & Orton, R. Day care regulation: The limits of licensing, *Young Children*, Vol. #2, 1980, September, 12-17.
- Collins, R. Child care and the states: The comparative licensing study, *Young Children*, Vol. #2, 1983, July, 3-11.

- Cohen, M. *New York City Indicator Checklist*, NYC Agency for Child Development, New York City, 1983.
- CWLA, Day Care Standards Revisions, *Child Welfare League of America*, New York City, 1983.
- Douglas, E. & Fiene, R. Making the almost impossible—possible: Evaluation of a human service (Ecological Paradigm for child care monitoring), paper presented at *National Association for the Education of Young Children*, November 1979, Atlanta, Georgia.
- Ferrar, H., Gleason, D. & Smith, B. *A State-of-the-Art Report on State Monitoring of Child Care Facilities*, Social Services Research Institute, Washington, D.C., 1980.
- Fiene, R., The two year old: Characteristics and management of his play, *Dimensions*, 1974, 2, 2, 46-58.
- Fiene, R., Current trends in research, *Children*, 1975, Spring, 15-18.
- Fiene, R. *Child Development Program Evaluation Scale*, Children's Services Monitoring Consortium, Washington, D.C., 1984.
- Fiene, R., Douglas, E., & Kroh, K. *Child Development Program Evaluation*, Bureau of Child Development, Harrisburg, Pa., 1977.
- Fiene, R., Douglas, E. & Nixon, M. *System Documentation for Pennsylvania's Instrument Based Program Monitoring Information System*, Children's Services Monitoring Consortium, Washington, D.C., 1983.
- Fiene, R. & Nixon, M. *Instrument Based Program Monitoring System*, Children's Services Monitoring Consortium, Washington, D.C., 1981.
- Fiene, R. & Nixon, M. *Indicator Checklist Methodology*, Children's Services Monitoring Consortium, Washington, D.C., 1983.
- Harms, T. & Clifford, R. *Early Childhood Environment Rating Scale*, Teachers College, Columbia University, New York, 1980.
- HEW, *FIDCR Report of Findings and Recommendations*, U.S. Department of Health Education and Welfare, Washington, D.C., 1978.
- Lazar, I. *Lasting Effects After Preschool*, U.S. Department of Health, Education and Welfare, Washington, D.C., 1979.
- Merrill, B. *West Virginia Indicator Checklist*, West Virginia Department of Social Services, Charleston, W.V., 1981.
- Morgan, G. *Regulation of Early Childhood Programs*, Washington, D.C.: Day Care and Child Development Council of America, Inc., 1971.
- Morgan, G. *Alternatives for Regulation of Family Day Care Homes for Children*, Washington, D.C.: Day Care and Child Development Council of America, Inc., 1974.
- Morgan, G. Regulating early childhood programs in the eighties, unpublished manuscript, 1980.
- NAEYC, Criteria for quality early childhood programs, *National Association for the Education of Young Children*, Washington, D.C., 1983.
- Peat, Marwick, & Mitchell, Assessment of the implementation of the indicator checklist for day care centers in West Virginia, *Children's Services Monitoring Consortium*, September, 1983.
- Poole, J. *Child Development Program Quality Review*, California State Department of Education, Sacramento, CA., 1982.
- Ruopp, R. *Children at the Center*, Cambridge, MA: Abt Books, 1979.
- Watson, S., Fiene, R., & Woods, L. *Kids MicroComputer Software*, Children's Services Monitoring Consortium, Washington, D.C., 1984.
- Zigler, E. & Gordon, E. *Day Care: Scientific and Social Policy Issues*, Boston, Massachusetts: Auburn House Publishing Company, 1982.

Theoretical Model For Computing Adult-Child Ratios in Day Care Centers



Dr. Richard Fiene

Dr. Richard Fiene here presents a theoretical model for computing adult-child ratios which takes both numbers (caregivers and children) and time (hours at the day care center) into account. Its purpose is to increase accuracy and save time in the computation of ratios to determine compliance with requirements, whether for FDCR or for licensing.

He is beginning an adaptation of this model for use in residential or twenty-four hour facilities.

He has confidence that his model can be put into State regulatory systems, using the expertise of program oriented licensers and statistics/research oriented technicians.

If you care to explore this presentation further, you may call Dr. Fiene at 717-787-2724. He is Director, Bureau of Information Systems, Pa.DPW, 1514 N. 2nd St., Harrisburg, Pa. 17102.

There has been much controversy over the Federal Day Care Requirements, in particular, the adult-child ratios. The majority of the discussion has revolved around the dichotomous points of view of the states and the federal role in enforcing the various standards. There is another issue that is equally important, which has been addressed only in a side-glance manner. Once it is decided what the ratios will be, how are we going to measure compliance with the ratios?

There have been various attempts at doing this - the most recent had been tried by Health, Education and Welfare (1977) and it does get at the required information. There is only one problem with it: it is rather time consuming. If a state or region has a great number of programs, it becomes almost an impossible task. Past methods have tried the direct approach of dividing the total number of children by the total number of teachers. This works, but does not give the overall day picture; therefore it is only good as a very gross measure.

The staff-child ratio question is a very critical item when it comes to monitoring of child development/child care facilities. However, it has eluded proper measurement because of inadequate or time-consuming measures. I am proposing a new theoretical model for computing adult-child ratios which is not time-consuming and gives accurate information in a very concise fashion.

With this new approach, all a day care monitor needs to do is ask six questions of the provider, and then put the data into a formula to find if the program is within compliance or not.

The six questions are as follows:

- (1) When does your first staff member (teaching) arrive?
- (2) When does your last staff member (teaching) leave?
- (3) Number of teaching staff?
- (4) Number of children present on your maximum enrollment day? Their ages and which staff members are assigned to each age group? (If there is vertical grouping).
- (5) When does your last child arrive?
- (6) When does your first child leave? (If vertical grouping, give breakdown according to age.)

After these questions are answered, then the day care monitor will compute the number of contact hours (CH) between staff and children using the following formula:

$$CH = \frac{NC(TH+TO)}{2} \quad \text{Formula 1}$$

In the formula, NC = total number of children present on the maximum enrollment day. TO = total number of hours the center is open. TH = total number of hours at full enrollment. CH = contact hours between staff and children in any type of caring arrangement.

After the CH is computed, the data are then put into another formula which will determine the relatively weighted contact hours for horizontal grouping (RWCH), or the relatively weighted contact hours range (RWCHR) for a vertically grouped program.

After the CH is computed, the data are then put into another formula which will determine the relatively weighted contact hours for horizontal grouping (RWCH), or the relatively weighted contact hours range (RWCHR) for a vertically grouped program.

$$RWCH = \left[\frac{NC(TH+TO)}{2} \right] \left[\frac{1}{TA} \right] \quad \text{Formula 2}$$

$$RWCHR = \left[\frac{\sum NC_i (TH_i + TO_i)}{2} \right] \left[\frac{1}{TA} \right] \quad \text{Formula 3}$$

$$NC_o = NC_p + NC_i + NC_s \quad \text{Formula 4}$$

$$TH_o = TH_p + TH_i + TH_s \quad \text{Formula 5}$$

$$TO_o = TO_p + TO_i + TO_s \quad \text{Formula 6}$$

In the formulae above, NC_o = total number of children on the maximum enrollment day. TO_o = total number of hours the center is open. TH_o = total number of hours at full enrollment. RWCH = relatively weighted contact hours - indicator of compliance for horizontally grouped programs. RWCHR = relatively weighted contact hours range - indicator of compliance for vertically grouped programs. TA = total number of teaching staff.

NC_i = total number of infants. NC_p = total number of preschoolers. NC_s = total number of school age children. TH_p = total number of hours at full enrollment with preschoolers. TH_i = total number of hours at full enrollment with infants-toddlers. TH_s = total number of hours at full enrollment with school-age children.

I = infant-toddlers. P = preschoolers. S = school-age children.

Once the RWCH or the RWCHR figures are computed, now we can find if the programs are within compliance by using the Table of Conversions for RWCH and RWCHR. (See Table 1).

This table is computed from an ideal where TO and TH both equal eight hours. In other words, all staff and children arrive and leave at the same time which is an ideal programmatic set-up. By using the Table of Conversions, it is relatively easy to compute if a program is within compliance.

I think a few examples will suffice: Example A: Day care monitor asks the six questions and gets:

- (1) 6:30 a.m.
- (2) 5:30 p.m.
- (3) six staff
- (4) 35 children - all four years old
- (5) 9:30 a.m.
- (6) 3:15 p.m.

Compute CH:

$$CH = \frac{NC(TH+TO)}{2} = \frac{35(5.45+11)}{2} = \frac{35(16.45)}{2} = 287.87$$

Compute RWCH: (because it is a horizontally grouped program).

$$RWCH = \left[\frac{NC(TH+TO)}{2} \right] \left[\frac{1}{TA} \right] = \left[\frac{35(5.45+11)}{2} \right] \left[\frac{1}{6} \right] = \left[\frac{35(16.45)}{2} \right] \left[\frac{1}{6} \right] = 287.87 \left[\frac{1}{6} \right] = 47.97$$

Now refer to the Table of Conversions. Look under NC = 35, CH = 280. Now look under RWCH in the P column. The score here is 56, which indicates that this program is well within compliance. In reading the Table of Conversions, if a program receives a score equal to or less than the score on the Table it will always be within compliance. If the program receives a score greater than the score on the Table for that particular category, then the program will always be out of compliance.

Example B: Day Care monitor asks the six questions and gets:

- (1) 6:00 a.m.
- (2) 6:00 p.m.
- (3) three staff
- (4) 15 children - five infants, five preschoolers, five school age children
- (5) 10:00 a.m.
- (6) 4:00 p.m.

Compute CH:

$$CH = \frac{NC(TH+TO)}{2} = \frac{15(6+2)}{2} = \frac{15(8)}{2} = 120$$

(Continued on page 11)

Theoretical Model For Computing Adult-Child Ratios In Day Care Centers (Continued from page 9)

Compute RWCHR: (because it is a vertically grouped program).

$$\begin{aligned}
 RWCHR &= \left[\frac{\sum NC_i (TH_i + TO_i)}{2} \right] \left[\sum \frac{1}{TA} \right] \\
 RWCHR &= \left[\frac{NC_1 (TH_1 + TO_1) + NC_2 (TH_2 + TO_2) + NC_3 (TH_3 + TO_3)}{2} \right] \left[\sum \frac{1}{TA} \right] \\
 &= \left[\frac{[5(6+12)] + [5(6+12)] + [5(6+12)]}{2} \right] \left[\frac{1}{3} \right] \\
 &= \left[\frac{90+90+90}{2} \right] \left[\frac{1}{3} \right] \\
 &= [135] \left[\frac{1}{3} \right] \\
 &= 45.00
 \end{aligned}$$

Now refer to the Table of Conversions. Look under NC = 15, CH = 120. Now look under RWCHR in the P column because we have an equal number of infants, preschoolers and school age children. The score here is 40-50 which indicates this program is well within compliance.

Example C: Day care monitor asks the six questions and gets:

- (1) 6:00 a.m.
- (2) 6:00 p.m.
- (3) Three staff
- (4) 20 children — all four years old
- (5) 7:00 a.m.
- (6) 5:00 p.m.

Compute CH:

$$\begin{aligned}
 CH &= \frac{NC (TH + TO)}{2} \\
 &= \frac{20 (10 + 12)}{2} \\
 &= \frac{20 (22)}{2} \\
 &= \frac{440}{2} \\
 &= 220
 \end{aligned}$$

Compute RWCH: (because it is horizontally grouped program).

$$\begin{aligned}
 RWCH &= \left[\frac{NC (TH + TO)}{2} \right] \left[\frac{1}{TA} \right] \\
 &= \left[\frac{20 (10 + 12)}{2} \right] \left[\frac{1}{3} \right] \\
 &= \left[\frac{20 (22)}{2} \right] \left[\frac{1}{3} \right] \\
 &= \left[\frac{440}{2} \right] \left[\frac{1}{3} \right] \\
 &= [220] \left[\frac{1}{3} \right] \\
 &= 73.33
 \end{aligned}$$

Now refer to the Table of Conversions. Look under NC = 20, CH = 160. Now look under RWCH in the P column. The score here is 53, which indicates that this program is well out of compliance.

The aspect of the above theoretical model is that it takes both time and numbers of staff into account. It is a simple one-shot mathematical calculation, and it can determine if a program is within compliance or not.

TC=8
TH=8

TABLE I
TABLE OF CONVERSIONS
FOR ADULT-CHILD RATIO

NC	CONTACT HOURS	RELATIVELY WEIGHTED CONTACT HOURS			RELATIVELY WEIGHTED CONTACT HOURS RANGE			NC	CONTACT HOURS	RELATIVELY WEIGHTED CONTACT HOURS			RELATIVELY WEIGHTED CONTACT HOURS RANGE		
		I	P	S	I	P	S			I	P	S	I	P	S
1	8	8	8	8+	8	8	8+	31	248	35	50	62+	42	43-49	50+
2	16	16	16	16+	16	16	16+	32	256	37	51	64+	44	45-57	58+
3	24	24	24	24+	24	24	24+	33	264	38	53	66+	45	46-59	60+
4	32	32	32	32+	32	32	32+	34	272	39	54	68+	46	47-60	61+
5	40	40	40	40+	40	40	40+	35	280	40	56	70+	48	49-56	57+
6	48	24	48	48+	36	37-48	49+	36	288	36	48	72+	42	43-60	61+
7	56	28	56	56+	42	43-56	57+	37	296	37	49	74+	43	44-61	62+
8	64	32	32	64+	32	33-48	49+	38	304	38	51	76+	44	45-57	58+
9	72	36	36	72+	36	37-54	55+	39	312	39	52	78+	54	46-59	60+
10	80	40	40	80+	40	41-60	61+	40	320	40	53	80+	46	47-60	61+
11	88	29	44	44+	37	38-44	45+	41	328	36	55	66+	45	46-51	52+
12	96	32	48	48+	40	41-48	49+	42	336	37	56	67+	46	47-52	53+
13	104	35	52	52+	43	44-52	53+	43	344	38	49	69+	43	44-54	55+
14	112	37	37	56+	37	37-46	47+	44	352	39	50	70+	44	45-55	56+
15	120	40	40	60+	40	40-50	51+	45	360	40	51	72+	45	46-56	57+
16	128	32	43	64+	37	38-53	54+	46	368	37	53	74+	45	46-56	57+
17	136	34	45	68+	39	40-56	57+	47	376	38	54	75+	46	47-57	58+
18	144	36	48	72+	42	43-55	56+	48	384	38	55	77+	46	47-58	59+
19	152	38	51	76+	44	45-57	58+	49	392	39	56	78+	47	48-59	60+
20	160	40	53	80+	46	47-60	61+	50	400	40	50	80+	45	46-61	62+
21	168	34	42	56+	38	39-46	47+	51	408	37	51	68+	44	45-53	54+
22	176	35	44	59+	39	40-47	48+	52	416	38	52	69+	45	46-54	55+
23	184	37	46	61+	41	42-49	50+	53	424	39	53	71+	46	47-56	57+
24	192	38	48	64+	43	44-53	54+	54	432	39	54	72+	46	47-56	57+
25	200	40	50	67+	45	46-54	55+	55	440	40	55	73+	47	48-57	58+
26	208	35	52	69+	43	44-51	52+	56	448	37	56	75+	46	47-56	57+
27	216	36	54	72+	45	46-55	56+	57	456	38	51	76+	44	45-57	58+
28	224	37	56	75+	46	47-56	57+	58	464	39	52	77+	45	46-58	59+
29	232	39	46	77+	42	43-53	54+	59	472	39	52	79+	45	46-59	60+
30	240	40	48	80+	44	45-61	62+	60	480	40	53	80+	46	47-60	61+