Measurement Tools and Systems

Chapter 11

Richard Fiene PhD
Karen E Kroh

◼ TABLE OF CONTENTS

■ 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.

<div align="center">

1   2   3   4   5   6   7   8

Low Risk   High Risk

</div>

---

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:

RULE VIOLATIONS	WEIGHTS
# 1	7
# 2	6
# 3	+ 8
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:

RULE VIOLATIONS	WEIGHTS
# 1	7
# 2	6
# 3	+ 8
Sum of Weights =	21

Final calculation:

$$
\begin{array}{r}
100 \\
\underline{-21} \\
79
\end{array}
$$

◼ **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 then 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.
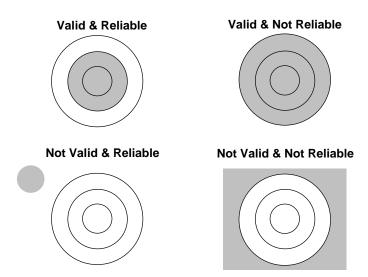
When there is not a close link between instrument development and rule formulation this only leads to headaches for licensing agencies. It may take years and not be evident until you get called into a court of law to defend your licensing system but it will happen.

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.
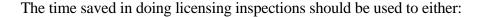
Reliability testing should be done methodologically and scientifically. Inter-rater reliability should be established for the tools/checklists that are to be used in the field by licensing field staff. This is a process that has been well documented in the psychological research. This has not been the 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.

NARA Licensing Curriculum, Chapter 11: *Measurement Tools and Systems*

■ 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.