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

Richard Fiene, Ph.D.

January 2021

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

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

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

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

a research psychologist whose total career (50 years) has been spent in early care and education, this has been the focus of my research.

The limitations of the above systems are numerous and, in some ways, are difficult to find solutions. In the past, these measurement systems have focused more on the descriptive aspects of data distributions rather than attempting to be predictive or inferential. The first major limitation of the data from regulatory compliance systems is the fact that the data distribution is markedly skewed. What does skew data mean? Most data distributions are normally distributed with very few occurrences at the extremes with the majority of the cases in the middle section of the measurement scale. IQ is an example of a normally distributed data distribution. In a skew data distribution, the majority of data are at one end of the data distribution, either at the positive end or the negative end of the distribution. With regulatory compliance data, it is at the positive end with the majority of facilities being in full or 100% compliance with the rules. Very few of the facilities are at the negative end of the distribution.

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

There is also a confounding factor in the regulatory compliance data distributions which has been termed the theory of regulatory compliance or the law of regulatory compliance diminishing returns. In this theory/law, when regulatory compliance data are compared to program quality data, a non-linear relationship occurs where either the facilities scoring at the substantial compliance level score better than the fully compliant facilities or there is a plateau effect and there is no significant difference between the two groups: substantial or fully compliant facilities when they are measured on a program quality scale. From a public policy stand point, this result really complicates how best to promulgate compliance with rules. This result has been found repeatedly in early care and education programs as well as in other human service delivery systems. It is conjectured that the same result will be found in any regulatory compliance system.

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

Let's turn our attention to potential solutions to the above limitations in dealing with regulatory compliance data.

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

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

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

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