

# Measurability Assessments

## House Bill 805, 2015 Session

Prepared March 22, 2016 by the Program Evaluation Division for the House Select Committee on Education Strategy and Practices

### 1. What are “Measurability Assessments”?

In 2015, the House passed the Committee Substitute for HB 805 to require the Governor’s Office of State Budget and Management (OSBM) to establish a system for assessing whether “a proposed or existing state program is or will be capable of reporting performance and return on investment.”<sup>1</sup> The bill “crossed over” but was not enacted, making it eligible for consideration by the 2016 short session. Measurability means that a program *should be capable* of producing records to support answers to such questions as:

- If a program is supposed to reduce a problem or increase well-being in North Carolina, how would the General Assembly know that its services (e.g. inspections, permits, workshops, etc.) are demonstrably linked to changes that can be measured? How was that link established? To what extent is the evidence valid and reliable?
- Does the program know and is it capable of reporting its direct and indirect costs?
- Can the program demonstrate how budget increases or decreases will affect results? (Scalability). Is there a point of diminishing return?
- Does the program have good management information systems for planning, organizing, accounting and reporting?

### 2. Why is independent measurability assessment necessary?

Measurability assessments are similar to credit ratings that are independent opinions of the risk that money lent or invested will not be repaid or fail to produce dividends. In state government settings, the taxpayer is the investor and the entity receiving taxpayer money is a proposed or existing program.

### 3. What will trigger a measurability assessment?

If enacted, the General Assembly by special appropriations provision or general law may require that an agency obtain a measurability assessment pursuant to H805 through OSBM.<sup>2</sup>

### 4. How will OSBM establish and administer the measurability assessment process?

HB 805 appropriates \$75,000 in non-recurring money to OSBM for a temporary contractor to design and assist with startup of the system.<sup>3</sup> The contractor will establish a method of assessing programs for the cost of assessments as provided by HB 805. The contractor will develop initial Requests for Qualifications to solicit potential assessors and assist OSBM in final selection. The contractor would also serve as a temporary resource to guide the first assessments. Once initiated, OSBM would assume complete responsibility for maintaining a pool of qualified assessors, supervising those assessors, collecting assessment fees from programs for assessment costs, and keeping assessment procedures up to date.

---

<sup>1</sup> New statute §143-E2 of the bill on page 1, lines 13-15.

<sup>2</sup> New statute §143-E2 of the bill on page 1, lines 13-15.

<sup>3</sup> See Section 2 of the bill on page 3, lines 13-15.

5. **What are qualifications for assessors and who will establish these qualifications? What assurance is there that assessors will produce credible and reliable assessment reports?**  
HB 805 requires OSBM in consultation with the Program Evaluation Division to set qualifications and criteria for selecting assessors through competitive recruitment. The bill specifies that OSBM cannot assign an assessor who has been employed by the entity (agency housing the program or making the proposal) within five years preceding the assessment.<sup>4</sup> Assessors should be experienced researchers familiar with quality assurance reviews, operations research, or research methodology. OSBM will more than likely pre-qualify assessors with the described technical skills from different policy areas of expertise (e.g. education, federal assistance, Medicaid, criminal justice, general government administration, or public transportation among others).
6. **Will the assessors duplicate the State Auditor, agency internal auditors, or the General Assembly's Program Evaluation Division (PED)?**  
No. The other organizations conduct detailed *post hoc* special or routine examinations using agency records, interviews, inspections, and analysis of data bases available during the time period examined. These audit organizations have often been hampered by the absence of measurable data that HB 805 addresses. Measurability assessments will determine *the capacity of a program to inform such examination* and where there may be gaps in systems and records. Assessments are not internal audits, program evaluations or financial audits.
7. **Would measurability assessments constitute "red tape" or additional "bureaucracy" that would delay a response to a crisis or impede an innovative solution?**  
No. H805 states that the General Assembly *may* require measurability assessments.<sup>5</sup> They would be imposed entirely at the General Assembly's discretion and the crisis or innovative character of the program may be a consideration.
8. **How will the measurability assessment program be funded?**  
A measurability assessment required by the General Assembly will be paid for through a fee assessment method established by OSBM. OSBM will assess the costs of the assessment consisting of the assessor's fee, travel expense if any, and a charge for OSBM overhead to the entity housing an existing program or proposing a program.<sup>6</sup> The Program Evaluation Division estimates that an assessment could range from \$5,000 to \$10,000 in cost depending upon the complexity of the program or proposal and the number of agencies or governmental subdivisions involved in implementation. An assessment should require from 50 to 100 billable hours. Assessors will be paid based upon a rate schedule established by OSBM.
9. **To what extent do current programs have the systems and data required by H805? Some of the criteria appear highly technical or difficult or costly to achieve. Will the bill force agencies to buy new technology or employ additional staff or consultants?**

---

<sup>4</sup> New statute § 143-E4 (a) of the bill, page 3, lines 2-4.

<sup>5</sup> New statute § 143-E2 (a) of the bill, page 1, lines 13-15.

<sup>6</sup> New statute § 143-E4 (c) of the bill, page 3, lines 5-8.

The bill does not mandate that agencies implement new systems or hire additional staff. The majority of existing programs will lack several elements because the criteria specified by H805 are “stretch” standards and have not been previously required by law. The purpose of H805 is to provide the General Assembly or agencies a tool to assess the existence of records and the **degree** of measurability of proposed or existing programs.

**10. Could a program or proposal “fail” a measurability assessment? Is this a “gotcha” exercise?**

No. After an assessment, the General Assembly will determine if the degree of measurability as reported by the assessor is satisfactory to justify the risk of new or continued funding. There will be no assignment of summary “pass” or “fail” scores based upon some grading system. While some of the measurability criteria are mutually exclusive attributes (“Have” or “Do not have,” “Yes or No, etc.), most call for a written opinion of the assessor as to the *degree* to which a standard has been attained and the circumstances affecting measurability.

**11. One of the criteria is if supporting evidence for a program was developed through “formative” or “summative” evaluations comparing “randomized” groups as part of an “experimental” or “quasi-experimental” design. What do these terms mean and why are they important?**

This aspect of measurability is the most technical, but is of utmost importance. Measurability assessors will be experts in research design capable of gauging the degree of validity of evidence purportedly supporting a program. Assessors will ask for any evidence grounding or underpinning a program and then determine the degree of validity of that evidence.

Evidence from a formative evaluation is for the purpose of establishing a new program. Evidence from a summative evaluation is developed after a program has been implemented. The difficulty in summative evaluation research is that denial of a legally-required ongoing service (such as an entitlement) from an existing public program for an experiment may be either illegal or unethical.

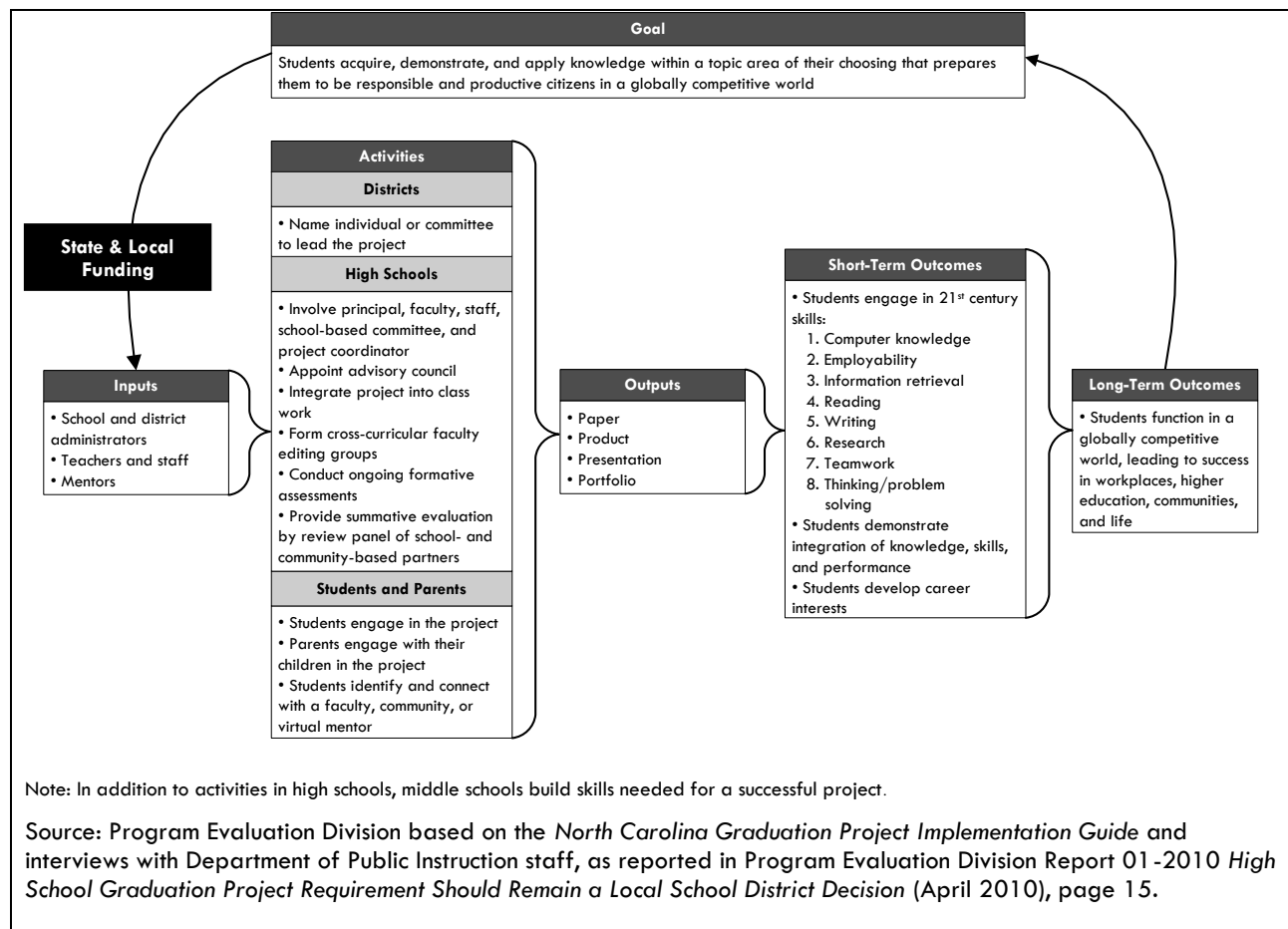
The “gold” standard for any research design is an **experimental design** where research subjects (those served by the program such as students, patients, clients, etc.) from the eligible population at large are randomly assigned to either the “control” group not subjected to program services or to the “intervention” group that is subjected. Differences are compared and measured for statistical significance. Results from experimental designs may be expected to be replicated and are largely generalizable to the entire population of subjects.

A **quasi-experimental** design is a “silver” standard used when randomization of subjects is impossible or impractical given time or budgetary constraints. Instead of randomizing subjects, researchers compare two groups of subjects (control and intervention) matched by exclusion/inclusion as equally as possible on variables such as gender, age, height, income, etc. that might influence behavior. Unlike experimental designs, results from quasi-experimental designs are less likely to be replicated if repeated, but are useful and widely applied in social science and educational research. However, they are less generalizable to entire population of subjects to the level of precision of experimental designs.

## 12. What is a “logic model” as called for by the bill?

A logic model is a diagram or flow chart that describes the goals of the existing or proposed program, funding, activities, units of output, short and long-term outcomes and how these components interact. A properly constructed logic model makes it easier to comprehend any concerted effort. H805 cites the *Logic Model Development Guide* by the W. K. Kellogg Foundation as a standard and resource.<sup>7</sup> In 2010, the Program Evaluation Division staff found that the Department of Public Instruction did not have a logic model for implementation of a state law requiring that a student must have completed a graduation project (similar to an Eagle Scout project) in order to graduate.<sup>8</sup> PED staff constructed a logic model for that program in 2010, which is shown below:

### Logic Model for the North Carolina High School Graduation Project Requirement as Operated in 2009-2011



<sup>7</sup> New statute § 143-E3 (b)(4) of the bill, page 1, lines 24-26.

<sup>8</sup> Subsequent to the PED report recommendation for repealing the statewide mandate, S.L. 2010-33 eliminated the mandate but allowed school districts to continue the graduation project requirement at local discretion. This allowed districts to avoid costs and reallocate up to \$750,000 in non-recurring and \$5,800,000 in recurring resources.

**13. Why are other criteria necessary for the assessment important?**

There are 14 assessment criteria within HB 805 as a proposed new statute § 143-E “Definition of measurability assessment.” In addition to the assessment of evidence under §§ (4) and the logic model criterion above, there are twelve other criteria that collectively represent characteristics of low risk, well-managed programs that are less likely to experience administrative failures, audit findings, benefit delays and backlogs, public complaints, media criticism, budgetary overruns, inefficiency or waste.