



# How to Write Great Performance Measures

Part of a Continuing Series on:

*Performance Monitoring and Program Evaluation*

Sponsored by NOAA's SEE Evaluation Committee

John Bortniak

Member, SEE Evaluation Committee

November 5, 2012



# Performance Measurement

***Ongoing monitoring and reporting of accomplishments, particularly progress toward strategic objectives***

## Uses of Performance Measures

- Strategic Plans
- Implementation Plans
- Annual Operating Plans
- Balanced Scorecard
- Quarterly / Annual Reviews
- Program Evaluations
- Program Management
- Personal Performance Plans



# *Milestone* versus *Measure*

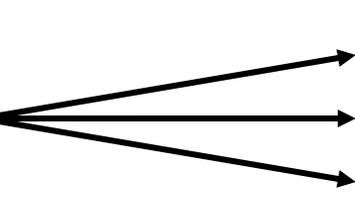
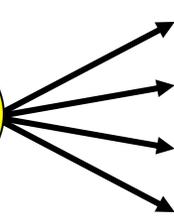
- A *Milestone* is a distinct activity planned for completion on a scheduled date. (*expires when completed*)
  - Launch a weather satellite on October 30<sup>th</sup>
  - Publish a proposed fishery rule to Federal Register, by June 1
  - Complete Marine Mammal survey using drones, by 4<sup>th</sup> quarter
  
- A *Performance Measure* is the monitoring of ongoing progress toward pre-established goals. (*quarterly targets*).
  - Increase tornado warning time. 4th quarter target: 30 minutes
  - Increase number of stream miles made accessible for migrating species. 1<sup>st</sup> quarter target: 120 miles
  - Square kilometers of seafloor mapped for coral habitat. 2<sup>nd</sup> quarter target: 400 Km<sup>2</sup>



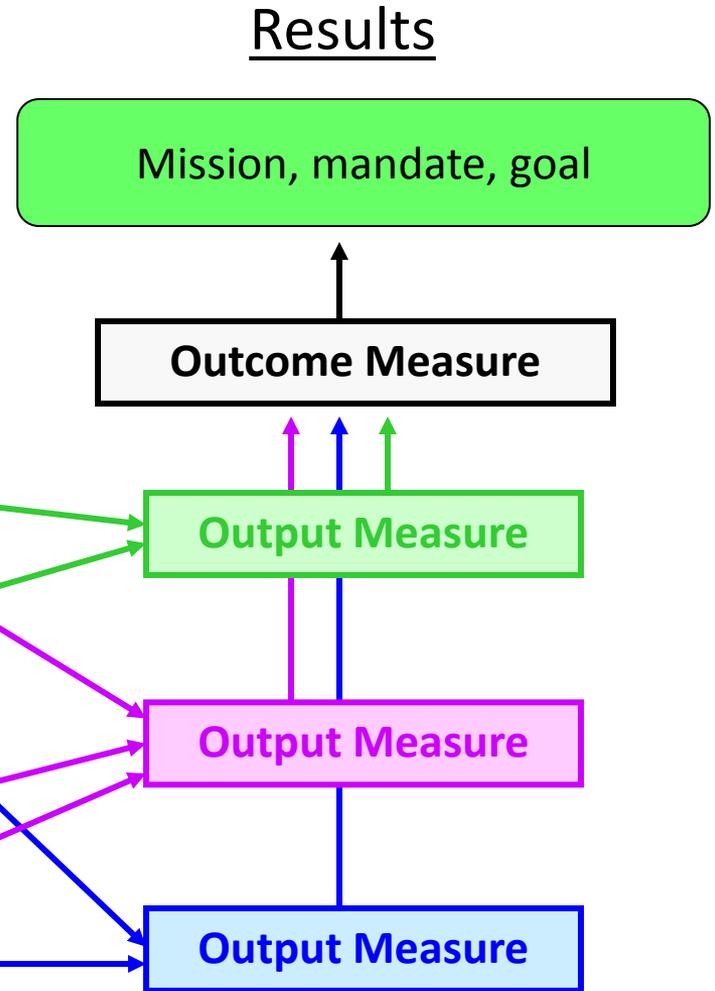
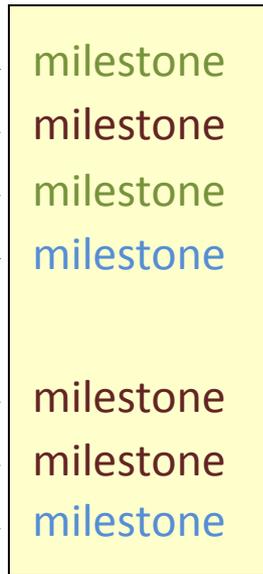
# Milestone-Measure Relationship (an AOP)

Performance Measure Targets should be supported by a set of Milestones intended to move you toward those Targets

## Infrastructure & Inputs



## Activities



Milestones are planned and implemented throughout the organization

Milestones support performance measures



# Components of a Measure

- All measures contain four components:

Indicator: what *change* is to be measured (the movement you want to see such as “Warning Time *Increased*”)

Unit of measure: how the indicator is to be quantified, either a number or % of: (minutes, stocks, acres, miles, etc.)

Baseline: starting reference level: how many (or %)

Target: desired performance: how many (or %)  
(All targets have dates associated with them)



# Example of a Measure

Decrease the sick patient's **fever back to a normal** temperature of **98.6 °F**

<u>Indicator:</u>	<b>fever back to normal</b>
<u>Unit of measure:</u>	<b>Temperature in °F</b>
<u>Baseline:</u>	102 °F, in the Morning
<u>Target:</u>	<b>98.6 °F, by Noon</b>



# Example of a Measure

Increase number of stream miles made accessible for migrating fish passage. 1<sup>st</sup> quarter target: 120 miles

Indicator: accessible for fish passage

Unit of measure: miles

Baseline: 4<sup>th</sup> Qtr FY 2012 = 40 (cumulative)\*

Target: 1<sup>st</sup> Qtr FY 2013 = 120 (cumulative)\*

\*(1<sup>st</sup> quarter is actually + 80 mile target)

- Whether a target (or actual) is cumulative, and from what starting point (annual or long-term) is a critical piece of information.



# Outcome Measures

- Define success based on mission and mandates:
  - *The true result - not the activities*
- Program's effect on the external environment
  - Number of Lives Saved
  - \$\$ in Property Damage Avoided
  - Number of Marine Mammal populations Recovered
- Answers the questions:
  - Why do we do this?
  - What are we ultimately trying to achieve?
- Define your Outcomes before the Outputs



# Trigger Questions for Outcome Measures:

- What is the mission of the program?
- What are its legislative mandates?
- Re-read your *Vision Statement*
  - If you were entirely successful... How would you know?
- What difference are we making in the world?
- Are there intermediate outcomes necessary for ultimate outcomes?



# Output Measures

- Focus on the quantity, quality, or timeliness of **products and services** delivered
- Program's internal work
- What most employees actually produce
  - Number of miles surveyed
  - Number of Terabytes of data archived
  - Number of Observer Days-at-Sea
- Logical linkage to achieving outcomes
- Useful to managers in day-to-day management



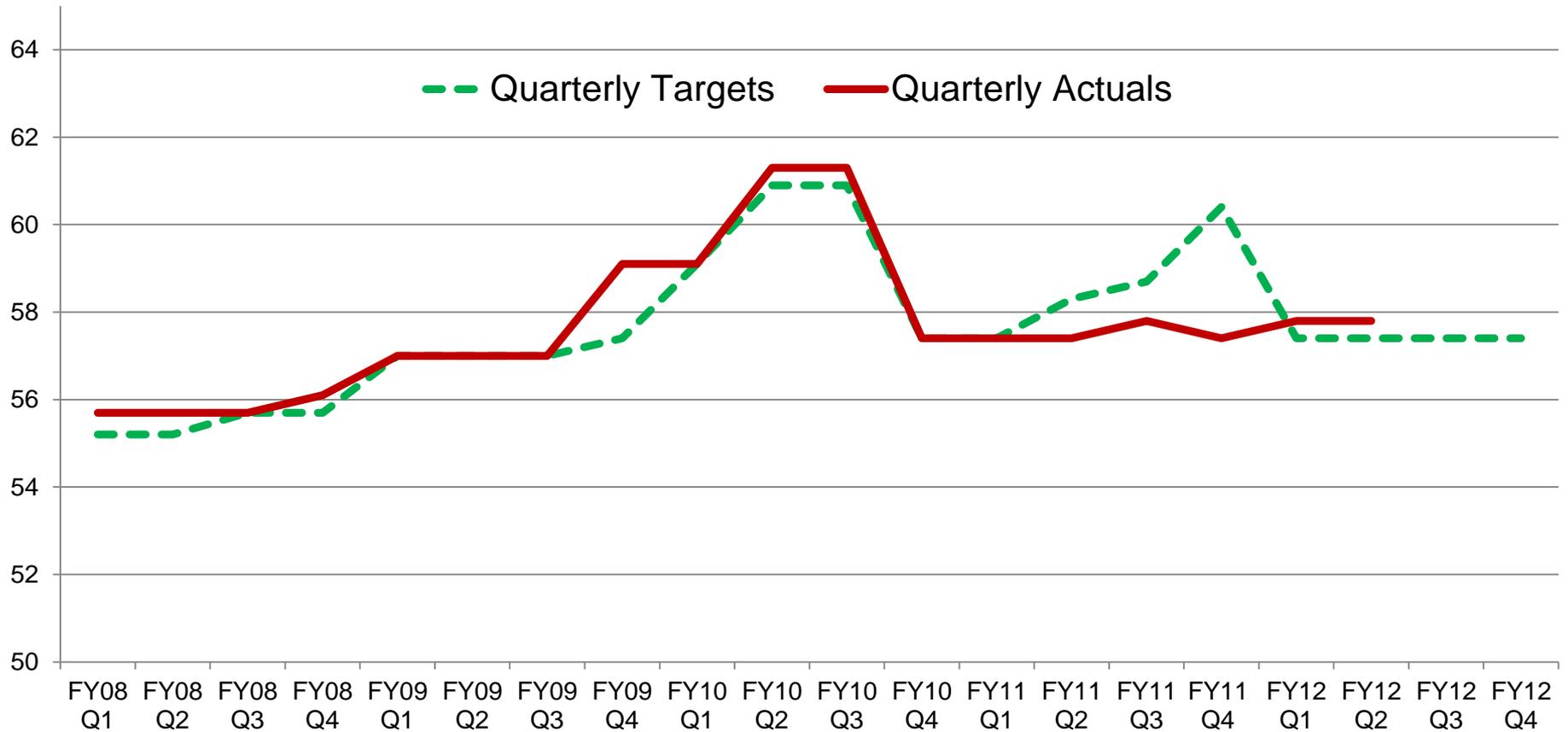
# Trigger Questions for Output Measures

- What conditions are necessary for the outcome to be achieved?
- What activities, products, or services are needed to create these conditions?
- What would an improved activity/product/service look like?
- Do we just want more, or also better or more timely?
- Can multiple activities/products/services roll into a single measure?



# An Output Measure

% of Fish Stocks with adequate stock assessments





# Efficiency Measures

- Expressed as ratio: 
$$\frac{\text{Units of Measure}}{\text{Cost to Produce}}$$

Example: 
$$\frac{\text{Miles surveyed}}{\$ \text{ million}}$$

- Can be either outcome or output measures



# The Performance Measurement Continuum

**Less Meaningful**

**Meaningfulness**

**More Meaningful**

**OUTPUTS**

**OUTCOMES**

Easier To Measure

More Budget Sensitivity

Less Activity-to-Results Lag

Easier to control

Harder To Measure

Less Budget Sensitivity

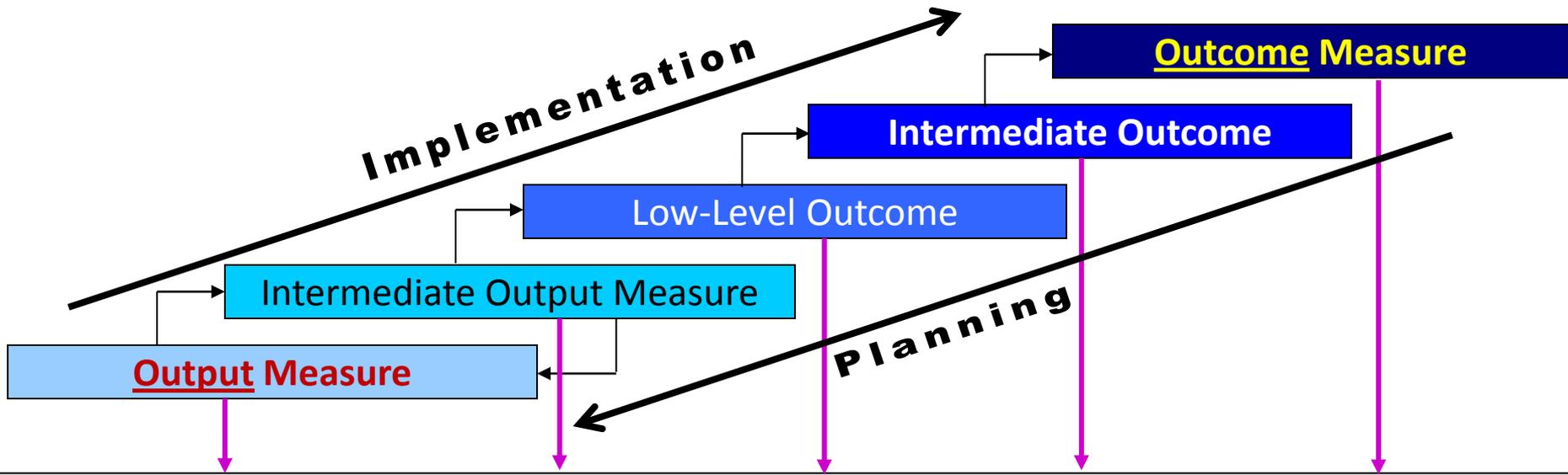
More Activity-to-Results Lag

Harder to control

**Measurability**

**Implementation**

**Planning**



The Measurability Timeline Scale (Data Update Frequency)

Quarterly

Annual

Decadal



# The Dilemma

No single performance measure can do everything:

## **Outcome** measures

- better demonstrate program benefits and success, but:
  - May take longer to measure
  - May be more difficult to measure
  - Don't fluctuate as much, or as quickly, as outputs

## **Output** measures

- may be less meaningful to the public, but:
  - Are more sensitive to funding changes
  - Show their effects more quickly
  - Are helpful to managers



# The Solution

A suite of measures is needed:

- Include both **outcomes** and **outputs** in your portfolio
- With measures closely linked, each having an effect on the next, with **outputs leading to outcomes.**
- Balance: *comprehensiveness and manageability*
  - **A dashboard - not a cockpit**
    - cover the whole program
    - but, keep the number of measures manageable.
- **Identify strategic points for measurement** in order to cover the key outcomes and outputs
- Include measures with **sensitivity to budget** changes



# Measures in the Budget

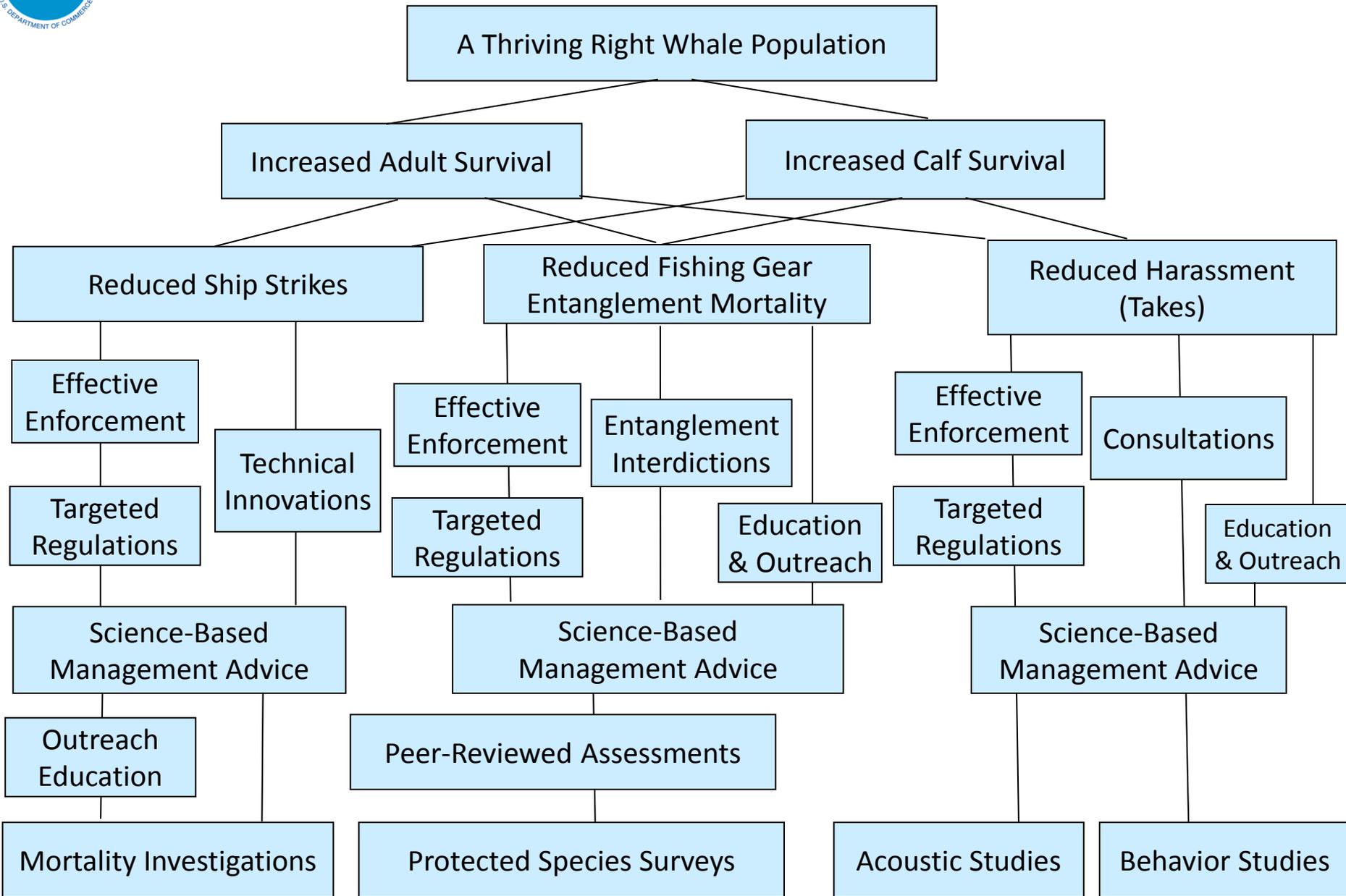
- Every budget increase must be supported by a link to performance

<b>Performance Measure:</b> <i>Acres of habitat restored for ocean, coastal, and Great lakes resources</i>	<b>FY 2010 Target</b>	<b>FY 2011 Target</b>	<b>FY 2012 Target</b>	<b>FY 2013 Target</b>	<b>FY 2014 Target</b>	<b>FY 2015 Target</b>
<b>With Increase</b>	<b>7,000</b>	<b>8,500</b>	<b>8,600</b>	<b>9,000</b>	<b>9,400</b>	<b>10,000</b>
<b>Without Increase</b>	<b>7,000</b>	<b>7,000</b>	<b>7,000</b>	<b>7,000</b>	<b>7,000</b>	<b>7,000</b>



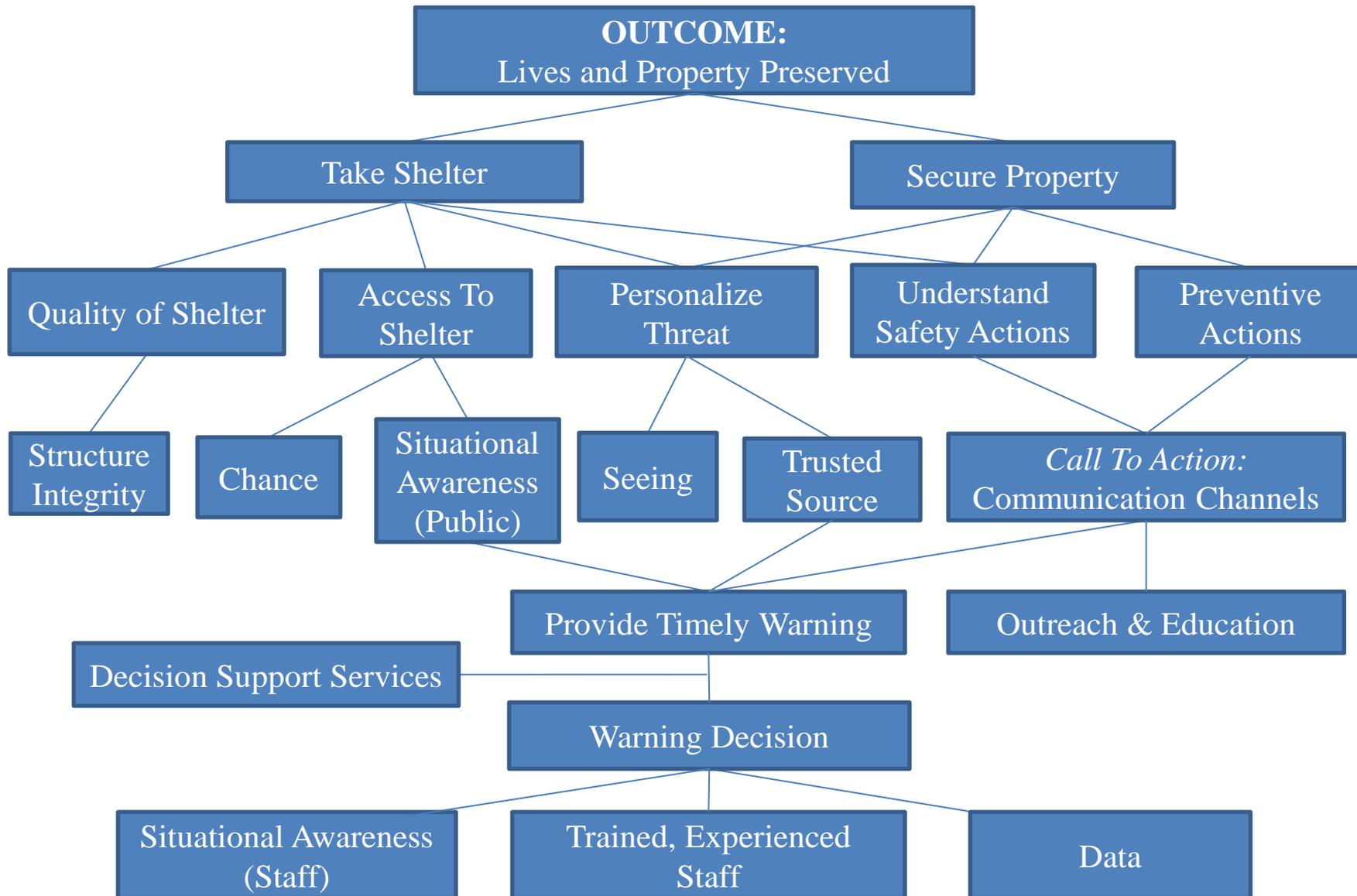


# Performance Measurement Linkages (LOGIC MODEL)





# Performance Measurement Linkages (LOGIC MODEL)





# Identifying Gaps in an Existing Suite of Measures

- Have you framed out a logic model?
- Are the Key *outcomes*, *conditions*, and *outputs* being measured?
- Do you have a distinct category of milestones (activities) that don't link to an existing measure?
  - Why are you doing them? To what end?
  - Maybe there's a missing measure?
- Do you have measures that demonstrate budget sensitivity?



# Quality Checklist for Outcome Measures

- Is the indicator truly meaningful?
- Is it the most important gauge of your success?
- Will it ring with the public and other customers?
- Can you actually measure it? Cost?
- Will the measure show at least gradual changes over a reasonable period of time?
- Will funding changes have at least a small demonstrable long-term effect?



# Quality Checklist for Output Measures

- Does an increase in the measure demonstrate improved functioning of the program?
- Is the measure a useful tool for managers?
- Is the measure focused on the output and not on what the output is supposed to achieve?
- Are your outputs or products clearly defined?
- Are the outputs within your control to achieve?
- Is there sensitivity to funding changes, in both short and long term?



# Quality Checklist for Efficiency Measures

Output      or      Outcome  
    \$\$                      \$\$

- Does the measure account for all costs associated with achieving the outcome or output?
- Which funding lines actually have an impact on the outcome or output?
- Does it account for delay between expenditure and result? (Is a time shift needed?)
- Is the measure useful to managers making resource allocation decisions?



# Business Rules

- Like a Cook Book, Business Rules document:
  - Definition of terms used in the measure
  - How to set targets, including sensitivity to budget changes
  - How to calculate actuals (counting method, formulas)
  - When do numbers become official? Who signs off? What's the trigger mechanism?
  - What's in the baseline (or denominator of a %)
  - Sources of data, and timing of data availability
  - Roles and responsibilities of HQ, field units, external entities



# Measurability

- Is the indicator measurable?
  - How precisely are the key terms defined?
  - Is it quantifiable?
- Have you chosen the right unit of measure?
  - Do you have influence over these units?
- Are data available?
  - Are the infrastructure and business processes in place to collect data?
  - If not, how difficult (or expensive) will it be to get data?
- How frequently is the data updated?
  - Will we see progress at least on an annual time scale?



# Relevance and Clarity

- Is the measure relevant to the mission?
  - Does it measure what *really counts*?
  - Do we have control to affect the outcome?
- Is the meaning clear?
  - The title of the measure should stand on its own without needing an explanatory paragraph to understand it
  - Avoid jargon, technical terms, and specialized acronyms
  - Avoid double negatives (reverse the decline of...)
  - Is an increasing or decreasing trend a good thing? Is that clear or confusing?



# Pitfalls

## Shifting baselines:

- Will elements be added or subtracted from the scope of the measure, changing numbers without changing performance (e.g. species listings/de-listings)?

## Number vs. percent:

- Percentages are preferred for measures where the universe is fixed and the goal is constant over time. When the 100% goal is unclear or changing, or the universe is frequently expanded or contracted (e.g. stocks listed under ESA), raw numbers are preferred.

## Threshold measures:

- Does the measure register performance only when a threshold is crossed, so that performance gains above and below the threshold will not register?



# Pitfalls (continued 2)

## Perverse incentives:

- Does a proxy unit of measure create a perverse incentive?

## Mixing apples and oranges:

- Does the measure add together things that sound similar but aren't?

## Handling unknowns:

- Will units about which there is currently no information potentially cause performance to drop when information becomes known?

## Illusion of stasis:

- Does the measure create the illusion of stasis in a dynamic situation?



# Pitfalls (continued 3)

## More efficiency through less quality:

- Does the measure track apparent efficiency increases achieved by reducing the quality of the outputs?

## Counting what's easy to count:

- Was the measure selected because it is truly the right way to gauge success of the program, or was it selected because it uses the most easily and readily available data?

## Simply measuring what we currently do:

- Output measures should be developed with careful analysis of what is needed for success, not simply “How should we measure the present outputs of this office/program?”

## Cumulative numbers not clearly identified:

- Be sure to clarify if numbers are cumulative and if so, from when?



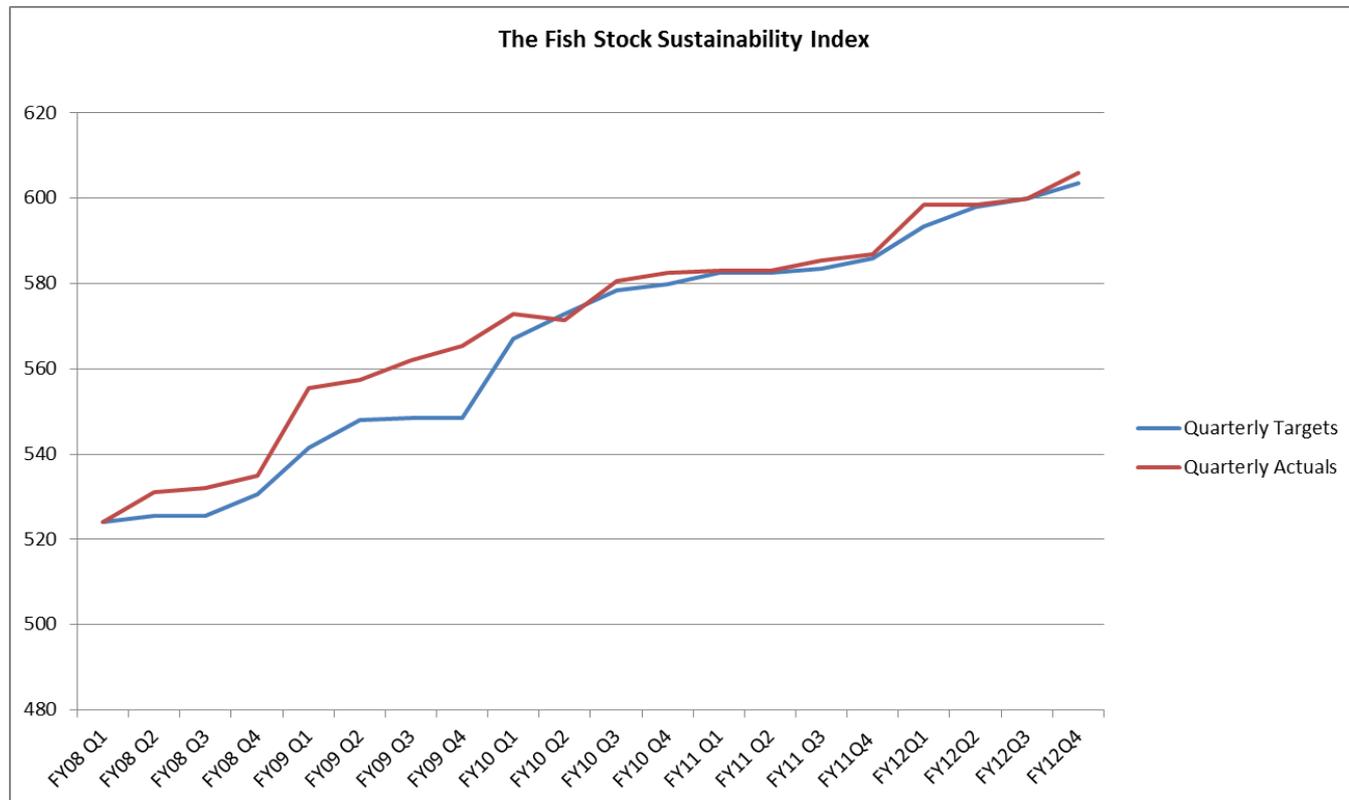
# Some Final Thoughts

1. Slow down, and using this guide, analyze your program or Office before you jump right into writing a measure
2. Take a step back, be willing to question assumptions, and ask seemingly stupid questions
3. Ask *What we should be measuring?*, not *How can we measure what we do?*
4. What if your program ceased to exist? What would be the difference? Could you measure that difference?
5. Asking why we do something moves the discussion toward outcomes
6. Asking how we do something moves it toward outputs
7. Good measures are often the result of thinking outside the box



# An “Index” Measure

- The Fish Stock Sustainability Index
  - Tracks status of 230 Fish Stocks
  - 4 possible points for each stock
  - 920 points is Total Success level





# Questions ?

Please fill out the online evaluation:

<http://goo.gl/Ovnzq>

Be sure to attend our **next presentation**

November 14, 2012:

***Overview of the DOC Performance Excellence Program***