

Program Evaluation & the Canadian Experience

April, 2010

Summary of Presentation

- overview of program evaluation
- overview of Canadian experience
- Auditor General's assessment of evaluation function
- sample evaluation (*One Tonne Challenge – to increase public awareness/participation in greenhouse gas emission reductions*)
- Lithuanian context

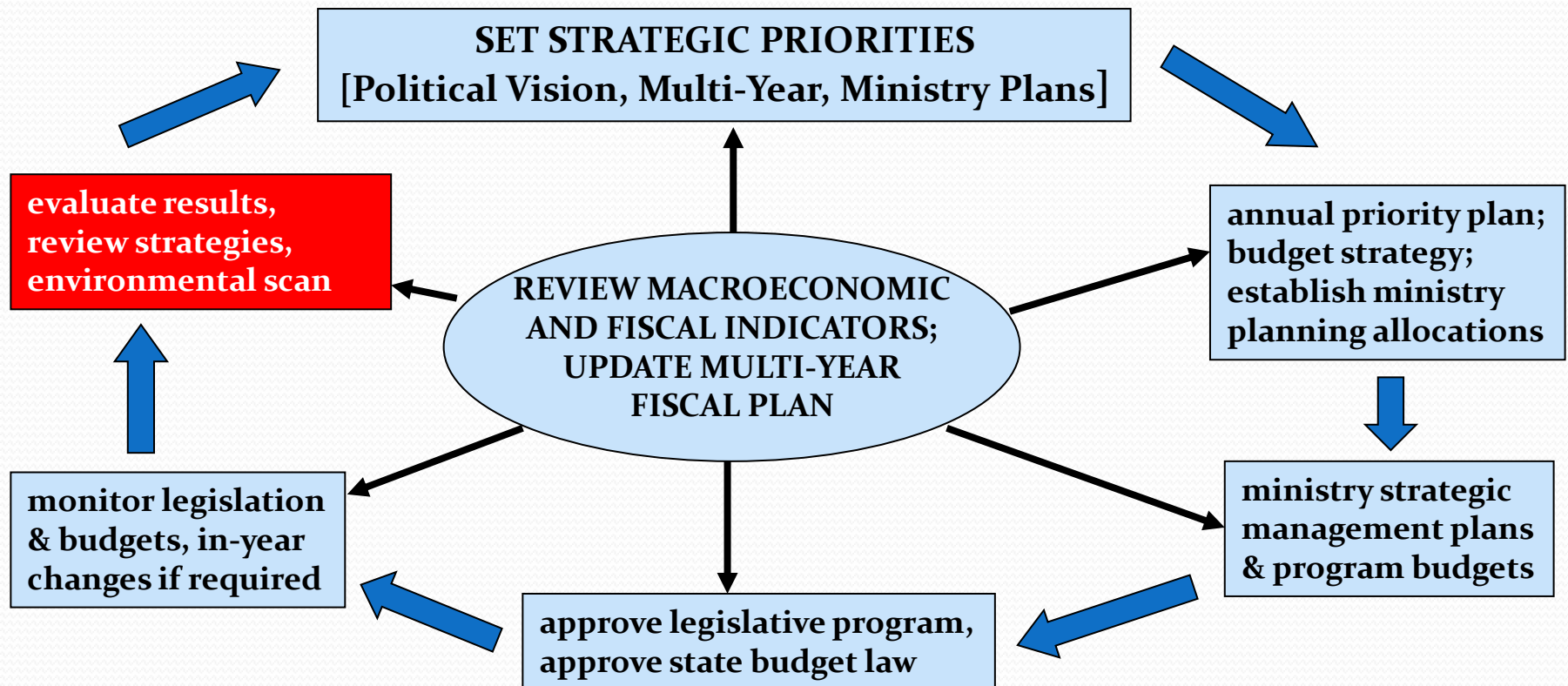
POINT OF DEPARTURE

Generally, Canada is considered as an advanced country in this area. However, even with its considerable advantages, its program evaluations are only partially successful. It will be important to consider why and to further consider how these challenges might be addressed in the Lithuanian context.

The Challenge of Program Evaluation

Evaluation: Important but Overlooked

- evaluation is a cornerstone of virtually every planning cycle developed in the last 40 years; nonetheless, it remains the **least developed** component in most countries



Why is this?

- **practical constraints**
 - skill constraints (*internal resources may lack necessary skill set*)
 - budget constraints (*external consultants are costly*)
- **technical constraints**
 - no baseline case may be available
 - relevant, reliable, historical data may not have been collected
 - programs may be poorly defined
- **administrative disincentives**
 - evaluation results likely viewed as judgement on senior management
 - implementing significant program changes is time-consuming, disruptive, and labour intensive
- **political disincentives**
 - politicians focus on implementing their promises (*Government Program*), not on improving ongoing programs (*unless this is a political commitment*)
 - poor evaluation results may provoke public criticism and provide opposition parties with ammunition to attack government (*except where government is newly elected*)

Why do it all?

- fundamental to **transparency**
- strengthens “**evidence-based**” decision-making
- increases **cost-effectiveness** (value for money)
- identifies **risks**; pre-empts major future problems
- establishes a culture of **continuous improvement** and quality assurance



Canada's Approach to Program Evaluation

Legal Framework for Evaluation

- first **Evaluation Policy** approved in 1977
- **Results-Based Management Policy** approved in 2001
- **Accountability Act** passed in 2006
- current Evaluation Policy approved in 2006; revised in 2009
- **5-year cycle** for all major programs
- requires establishment of an objective **evaluation function** within each ministry and major agency
- **linked to strategic reviews** of expenditure management system
- evaluation policy itself must be evaluated every 5 years

Establishing an Evaluation Function

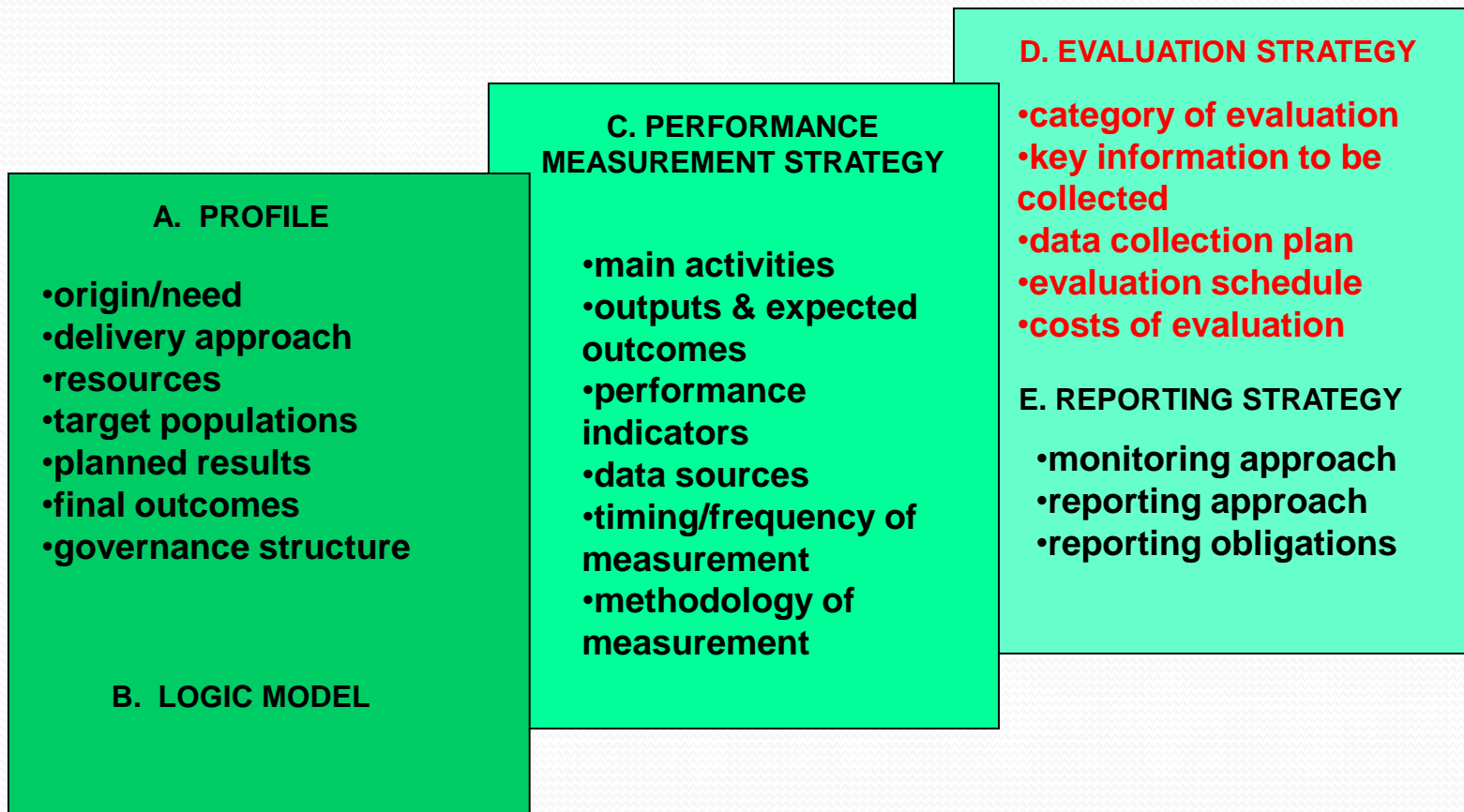
- Treasury Board established a **Centre of Excellence** in Evaluation (**12 staff**)
- evaluation departments vary in **size: 7 to 54**
- total of **404 evaluation staff** across government
- high **reliance on consultants** (90% of evaluations)
- ministry **evaluation budgets** range from \$1 million to \$14 million
- **annual evaluation plans** prepared and overseen by evaluation committee

Critical Success Factors

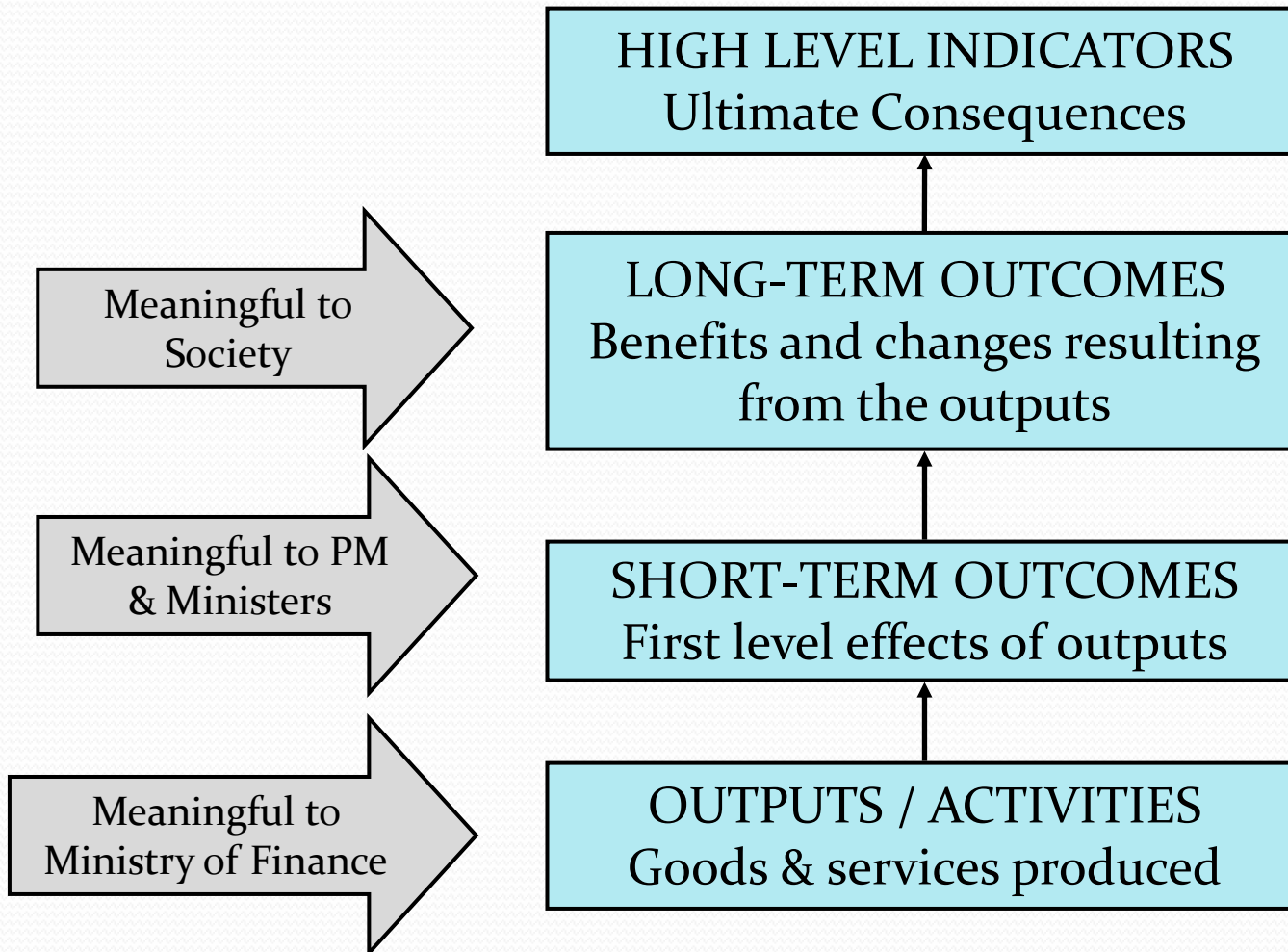
- program evaluation is viewed as a **core element of the planning cycle** and accountability framework
- preconditions for effective program evaluation include:
 - clear program structure exists
 - sound performance measures
 - clear targets
 - credible data
- the Canadian system does meet these preconditions
- in the absence of these, evaluations would be largely anecdotal and questions of validity and objectivity would arise

Approving New/Revised Programs

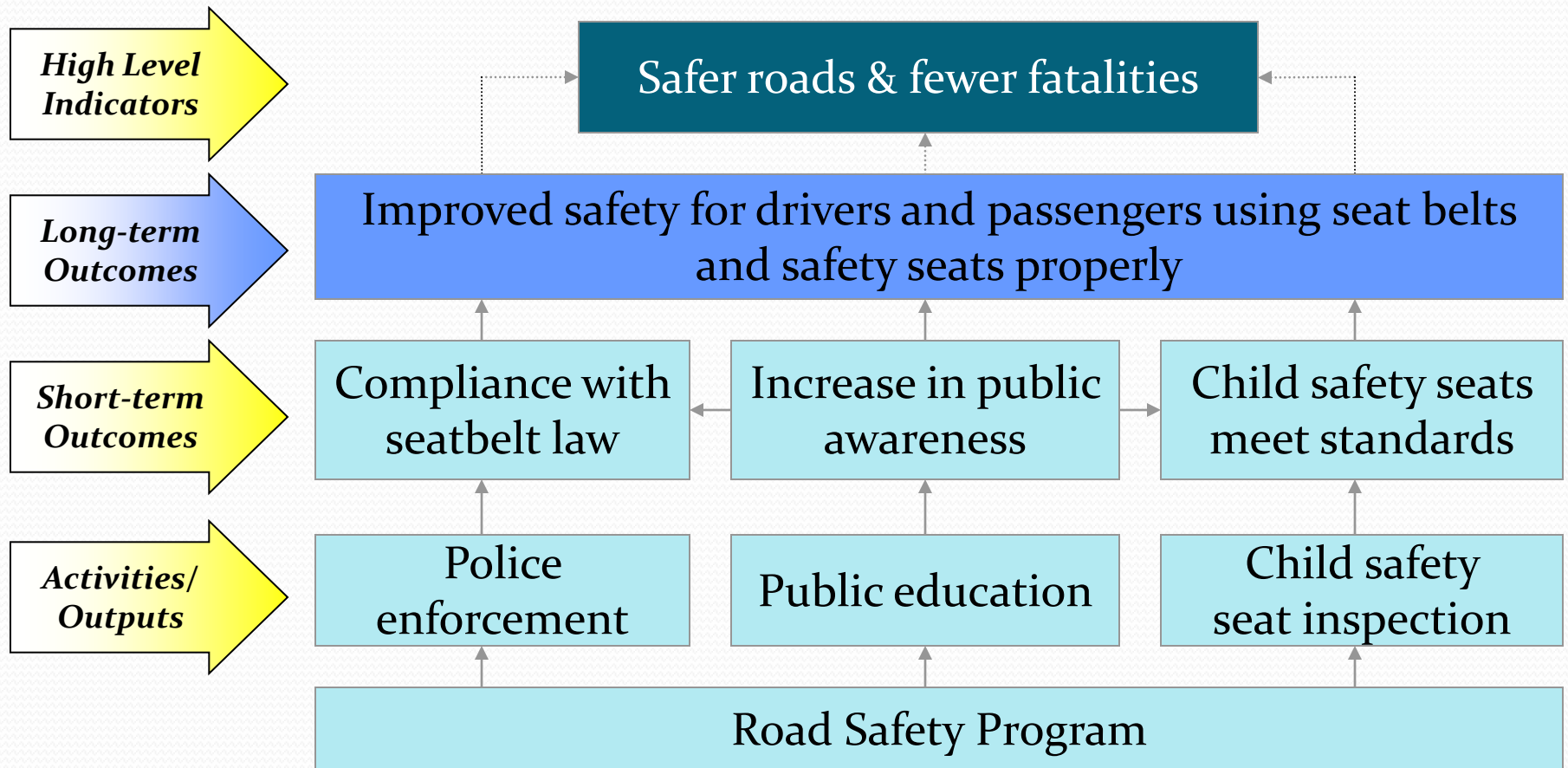
Results-Based Management & Accountability Framework



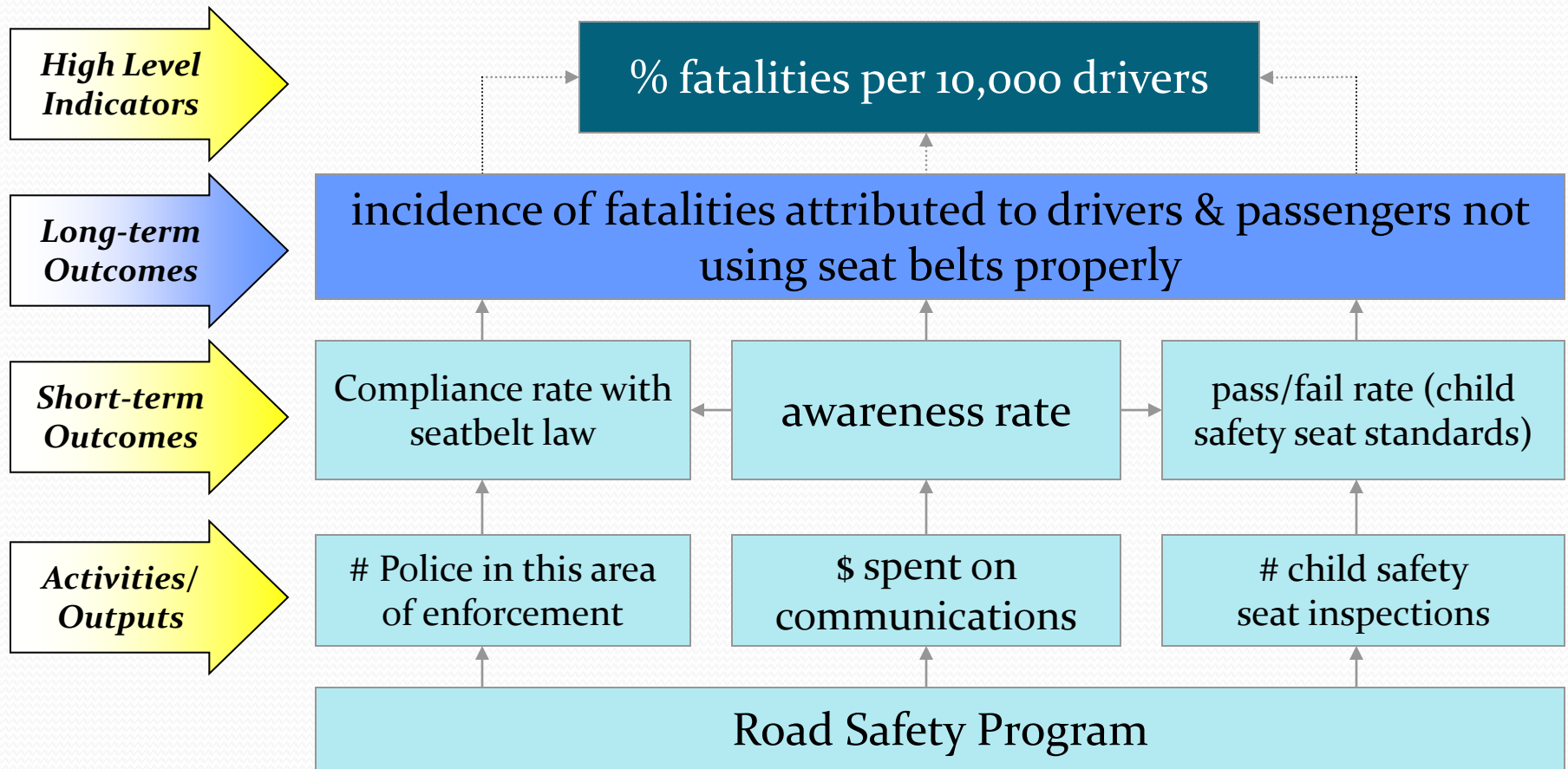
Logic Model for Determining Performance Measures



Developing Measures Using Logic Model [Road Safety Program]



Logic Model Applied to Road Safety



Linked to Budget

PROGRAM X					
YEAR ONE		YEAR TWO		YEAR THREE	
staff level	budget	staff level	budget	staff level	budget
X	Y	X	Y	X	Y

EXPECTED RESULT	PERF. INDICATOR	PERF. TARGET (year 1)
X	X	X
Y	Y	Y
Z	Z	Z

Planned Actions to Achieve Results

Benefits for Canadians

Linked to Strategic Reviews

STRATEGIC REVIEWS

- high-level, high profile
- may review entire ministry or major horizontal policy (groups of programs)
- may recommend significant structural or policy changes
- proposals can be controversial



PROGRAM REVIEWS

- low to medium-level and profile
- reviews one program (*although related program reviews may be linked*)
- typically recommends process and delivery improvements; rarely challenges purpose of program
- proposals rarely controversial

External Program Evaluations

- external program evaluations are conducted by **Auditor General** independently of government
- published in annual report
- Auditor General regularly **assesses quality of evaluations by Government**

Auditor General Reports

- reveals that the evaluation function has traditionally been problematic
- **1983 report:** “few high quality evaluations are being conducted”
- **1996 report:** “little progress on evaluation of effectiveness”, or study of major programs

Auditor General 2009 Assessment

- improvements noted (*evaluation function; policy, annual plans*)
- reviewed evaluations of 6 ministries
- evaluation plans were in place, but
 - **low proportion of expenditures** covered (*avoiding big programs*)
 - **reliant on interviews** rather than hard data (*poor data collection*)
 - 70% of evaluations **did not adequately assess “effectiveness”** (*which requires credible performance measures and targets on expected policy outcomes*)
 - **recruitment problems** (*departments understaffed*)

Sample Evaluation: One-Tonne Challenge

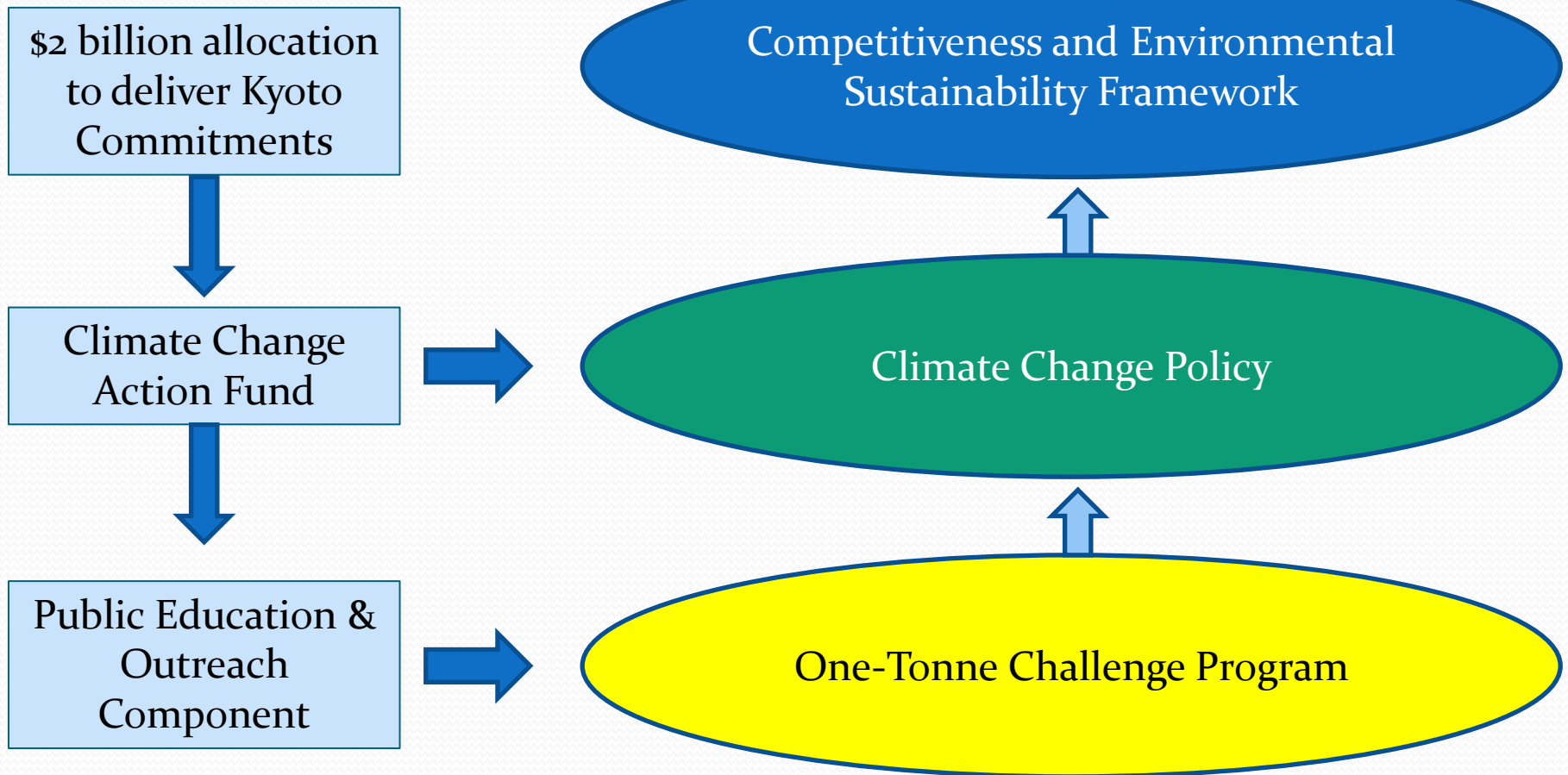
One-Tonne Challenge: Context

- part of **Kyoto** commitments
- at the time the program was launched (March 2004), individual Canadians produced an average of 5 tonnes GHG emissions per person annually
- OTC intended to serve as a catalyst for achieving a **20% reduction** in this amount (*i.e., reduce by one tonne*)
- OTC decided to use **public education** (*national marketing campaign*) and **outreach** (*establish partnerships*) as one approach for achieving this goal

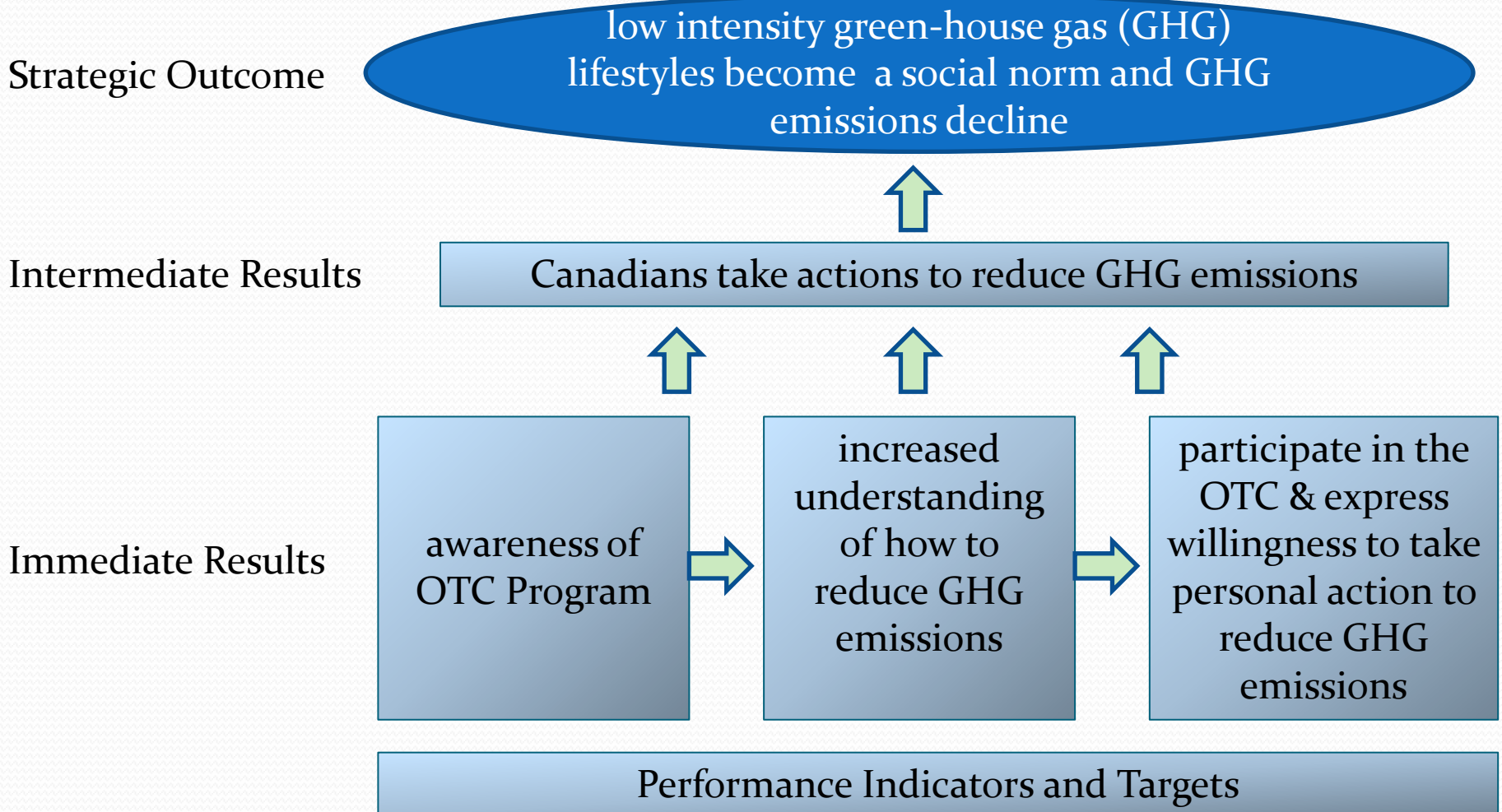
Note: one tonne = 1,000 kilograms

OTC: Policy & Funding Framework

FUNDING



Logic Model of One-Tonne Challenge (OTC) Program



Evaluation Overall Design

Evaluation Issue	Evaluation Question
1. relevance	Does the OTC continue to be aligned with the government's and ministries' priorities?
2. success	Is the OTC on track to meet its intended outcomes?
3. cost effectiveness	Are the most appropriate and efficient means being used to achieve outcomes?
4. design and delivery	Is the OTC being designed and delivered in the best possible way

evaluation covered first two years of program: March 2004-06

Evaluation Detailed Design

Issue	Indicator
1. relevance	1. alignment with approved government/ministry priorities and policies
2. success	1. confirm performance data collection 2. evidence that information used to influence management decisions 3. % Canadians aware of OTC 4. % Canadians aware of ways to reduce emissions as a result of OTC 5. % Canadians willing to take action in OTC 6. % Canadians taking action in OTC (compare to yes/no in Q5) 7. % change in GHG emissions attributable to OTC
3. cost effectiveness	1. budget allocations per program component 2. targets/objectives achieved
4. design and delivery	1. comparison with other countries' approaches/outcomes 2. synergy with similar programs delivered by Canadian provinces 3. actions implemented against program design 4. reach per program component 5. partner/stakeholder satisfaction 6. Canadian public opinion regarding OTC 7. performance comparison with other social marketing campaigns

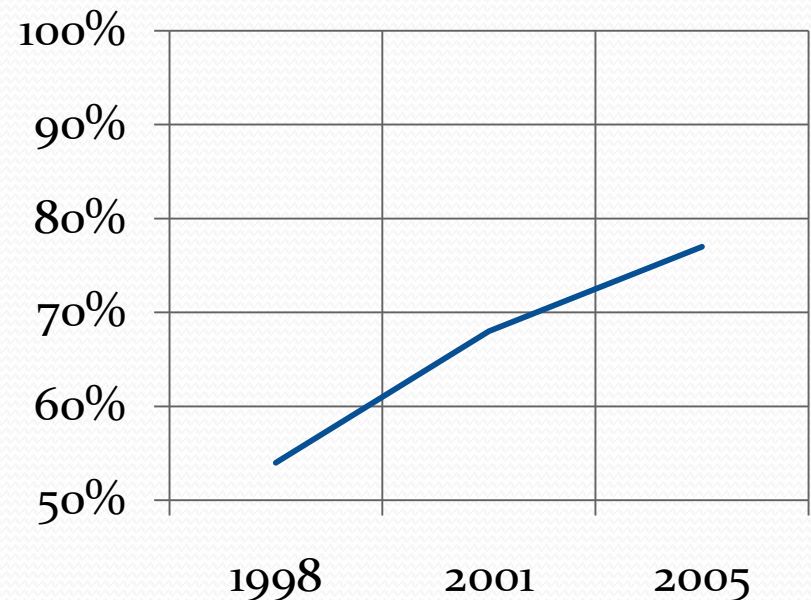
Evaluation Methodology

- **documentation review**
- **interviews** with officials and stakeholders
(administered by polling firm)
- **survey review** *(administered annually by polling firm from 2003)*
- **program mapping** *(to understand linkages)*

Findings re Relevance

- OTC remains relevant
 - remains government priority
 - individual citizens believe they can do something
 - information gap cited as main barrier to action
- but challenges remain
 - only 20% Canadian “extremely concerned about climate change
 - majority citizens (68%) do not actively participate

% Canadians who felt they could do something about climate change



Findings re Success: Awareness

- in 2004 (*just after program launch*), 6% of Canadians had heard of OTC)
- in 2005, increased to 51%
- 70% attributed awareness to a national television campaign
- of those who were aware of OTC, 78% understood the purpose of the program

Findings re Success: Participation

- of those who had heard of OTC, 54% claimed to be participating (up from 49% in 2004)
- represents 32% of population (up from 27% in 2004)
- of participants, 58% said that it was not difficult
- of non-participants, 54% indicated that they may participate in future years

Findings re Success: Non-Participants

- of non-participants, 64% claimed it would be difficult to meet 20% reduction target
- profile of non-participants indicates that highest GSG emitters are most resistant or sceptical about OTC

Group	% citing difficulty
higher income	73%
house owners	69%
car owners	65%

Findings re Success: Barriers

- main barriers to meeting GHG emission reduction goal included lack of information and inconvenience
- technical measurement and attribution problems cited as reason for not evaluating impact of OTC on actual GHG emission levels

Reason	% See as Barrier
lack of information	42%
inconvenience	33%
indifference	28%
perceived cost	26%

Findings re Design & Delivery

- overall **design influenced** by lessons learned from past programs (*used partnerships rather than individual projects*)
- generally felt to be **well-designed and delivered**, but two key areas to improve included:
 - poor coordination of youth and education components
 - poor integration with provincial government programs (e.g., energy conservation)
- **spin-off benefits** for related programs very positive (*OTC contributed to higher take-up*)

Findings re Cost Effectiveness

- because of diffuse nature of climate change issue (*precise consequences less clear*), will take **longer to generate buy-in** than comparable programs with more direct consequences (*e.g., anti-smoking*)
- OECD studies suggest that **public education programs need to be accompanied** by regulatory and economic instruments to move citizens from awareness to action

Conclusions

- program remains relevant
- design & delivery generally efficient; select improvements could be made
- awareness goals likely to be reached
- participation goals more of a challenge as highest emitters indicate least willingness to participate
- not realistic/feasible to link program directly to actual GHG emission reduction targets
- public education approach needs to be supplemented
- cost-effectiveness would be enhanced if accompanied by more direct regulatory and economic instruments
- measures to exploit economic growth potential should be considered (*e.g., supporting expansion of green industries*)

Reading Between the Lines

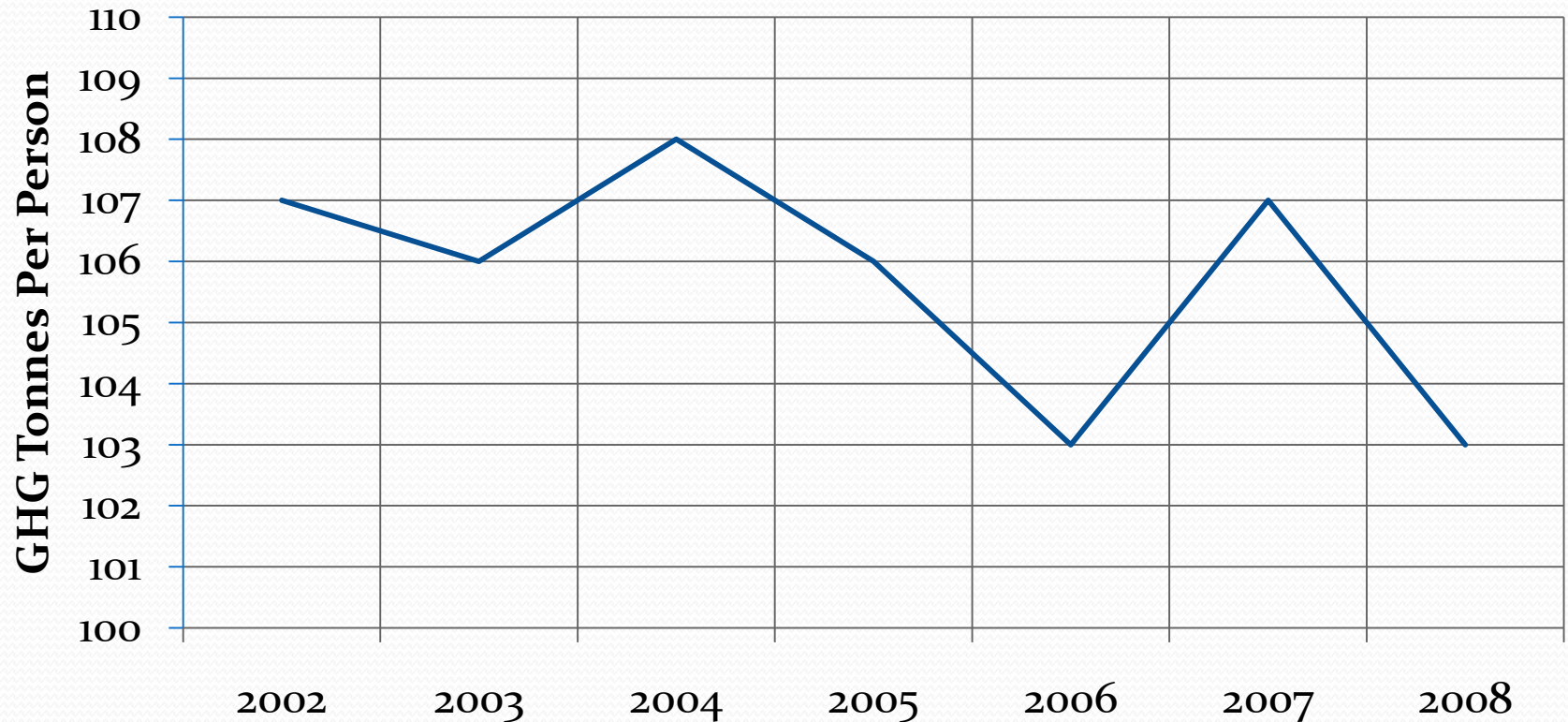
- only tangible evidence of success is the evidence of **increased public awareness** that the OTC program exists
- by contrast, the more important objectives regarding **program participation** and **concrete results** (*reduction in GHG emissions*) had not indicated significant progress
- likely known at the time that the target (20% reduction) was **more aspirational than realistic**
- given experience with similar programs (*anti-smoking*) and abstract nature of climate change issue, these results were not surprising

What Happened Next?

- OTC program put in place by Liberal Government in 2003
- Conservatives won 2006 election (*campaigning on promise that economic growth would be prioritized over meeting Kyoto commitments*)
- OTC **program cancelled**
- Canada remains 2nd worst (*behind Australia*) GHG emitter of OECD countries

Canada's GHG Emissions

Per Capita GHG Emissions



Conclusions

- OTC was a highly visible program launched by the Liberal government to address a **high-profile commitment** by the Prime Minister (join Kyoto)
- severe criticism of the OTC would have been **perceived as direct criticism** of the Minister and the Prime Minister
- the evaluation itself **strikes a balance between tone** (*which is neutral to positive*) **and content** (*which is neutral to negative and expresses doubt about the public education approach as a driver for achieving the Kyoto commitments*)
- overall, the evaluation **retains objectivity** (*facts were not hidden*) without creating political problems (*didn't say that OTC was a failure*)

Politics and Evidence

- evaluation called attention to the need to supplement the public education (*with economic and regulatory measures*) approach as public education alone was not achieving the desired pace of change
- if the current Government had not been defeated, they would likely have accepted this evidence-based recommendation
- in cancelling the program, the new Government could also defend its decision as being evidence-based
 - public education alone was not delivering the desired results
 - the Government had no intention of introducing economic or regulatory measures to supplement the public education approach
- the lesson: there is **rarely a single interpretation of evidence**

Best Practices in the Lithuanian Context

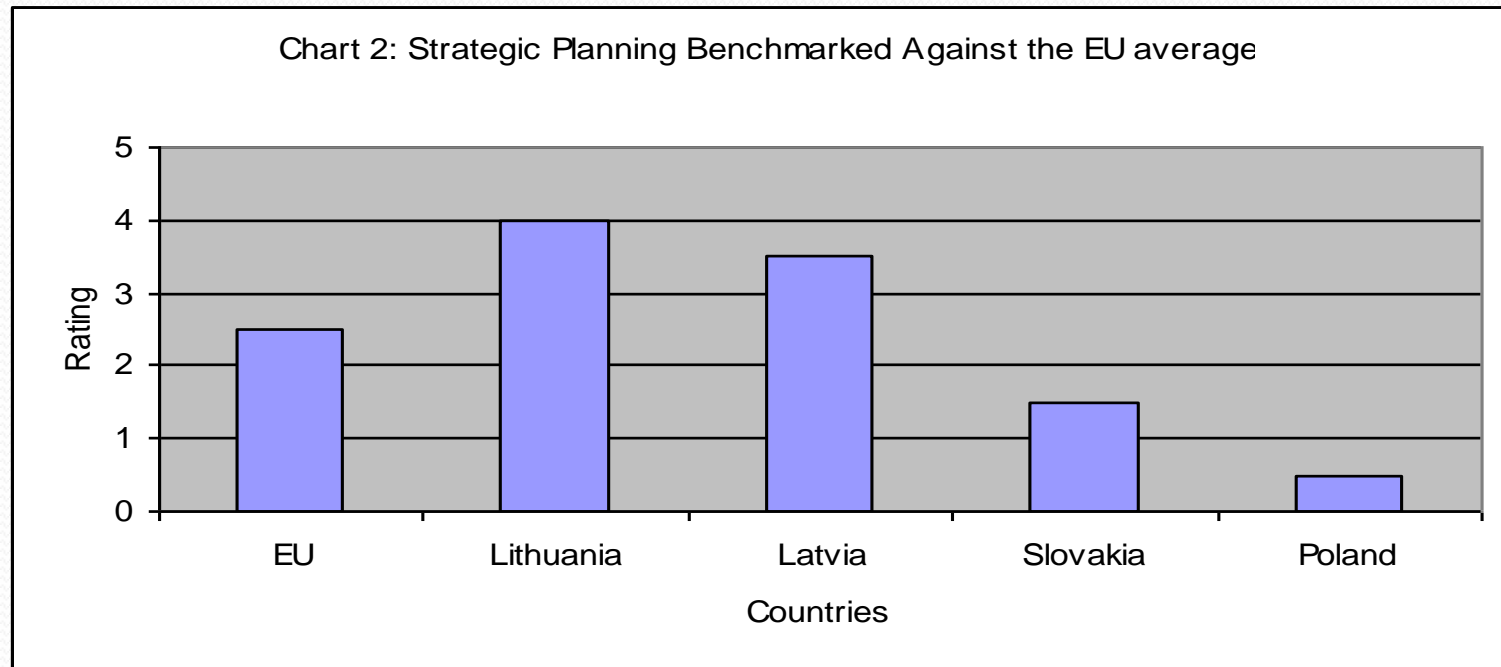
Lithuanian Context

- what issues/lessons learned from the Canadian context could be relevant to Lithuania?
- what approach should Lithuania take to address these issues?
- what other measures should Lithuania consider?
- what are Lithuania's relative strengths/weaknesses with respect to delivering program evaluations?
- what measures should be considered to introduce and continuously improve the evaluation function?

Lithuania: Starting Point

- several **natural advantages** (*small country; unitary system*)
- administrative experience is stronger than in most comparable countries
- **long-established basic systems** in place
 - strategic planning
 - program budgeting
 - results reporting
- **supportive political environment**
- **recent fiscal crisis** creates incentives to look for efficiencies and savings

Lithuania has Strategic Planning Capacity



WORLD BANK STUDY: 2006

PE must be Integrated

- ✓ incorporated within **strategic planning methodology**
- ✓ incorporated within **annual government work plan**
- ✓ informs **policy decisions**
- ✓ informs **expenditure decisions**
- ✓ coordinated with related exercises (*e.g., strategic or expenditure reviews*)
- ✓ takes **budget cycle** into account
- ✓ takes **political cycle** into account

Ensure Reasonable Autonomy

- ✓ **policy/regulations establish autonomy** of those responsible for conducting the evaluation
- ✓ ideally, **political level should reinforce** commitment to this principle
- ✓ evaluations should **focus on “evidence”**, not on opinion
- ✓ **avoid highly political issues** (*special or external reviews –supreme audit institution- can address these*)
- ✓ results should be used to improve programs **not to punish program managers** (*except in extreme cases*)

Clear Accountability

- ✓ results should be **presented to decision makers**; clear direction provided
- ✓ part of a **comprehensive RBM system**
- ✓ includes **logic model**; results chain
- ✓ **roles/responsibilities** clearly defined
 - ✓ decision makers
 - ✓ central institutions
 - ✓ ministries/agencies
 - ✓ implementation responsibilities
- ✓ reports are published (for parliament, public)
- ✓ “evaluation” is periodically evaluated

PE is built in from the beginning

- when creating a new, or reviewing an existing program (or set of programs), ensure that:
 - **program objectives** are clear
 - good quality **outputs** are identified
 - **resources can be attributed** to those outputs
 - **reliable, obtainable data** will be captured and reported
 - a **performance baseline** is established
- if solid data cannot be captured (*e.g., for programs with abstract or shifting outputs, such as those in foreign affairs*), determine **alternative ways** of establishing baselines and assessing performance (*e.g., customer surveys; comparative assessments with other countries*)
- incorporate a **PE Plan** within the documentation establishing (or restructuring) the program (or sub-program)

Plan Carefully

- ✓ put effort into a **well-crafted evaluation plan** (*or terms of reference if contracting to consultants*)
- ✓ fully **understand what data is available**
- ✓ **stay involved** (*if contracting to consultants*)
 - ✓ ensure you get the promised consultants
 - ✓ require early and interim products (*e.g., work plan; inception report; initial findings*)
 - ✓ review products at key points with evaluation oversight committee

Start Slowly, Choose Wisely

- do **not overload system**; build up to a cycle (*e.g., all programs assessed every five-years*) gradually
- avoid highly **complex programs**
- avoid politically **contentious programs**
- but, initial evaluations should **generate political interest**
- pick programs that can be **easily evaluated**
 - baseline exists
 - historical data exists
 - outputs are concrete
 - results would be valued by political level

Remember the Politics

- although PE is largely a technical exercise, important to **take political environment into account**
- avoid **ideological debates**
- avoid **politically controversial issues** that are currently in the media
- identify opportunities where the evaluation results may **support delivery of government priorities**

Know When to Contract Out

- typically, **civil servants can develop evaluation skills** in a wide range of techniques and methodologies
 - questionnaire development and application
 - basic costing of outputs
 - customer surveys
 - simple cost-benefit assessment
- however, there may be **some skills that need to be obtained from outside expertise**
 - in-depth cost-benefit analysis
 - focus groups
 - public opinion polling
 - specialized knowledge (e.g., major public investments)
- may also be occasions when contracting out is required to ensure that the evaluation is **perceived as being objective**
- **time commitments** may also determine if consultant are needed

Build an Evaluation Network

- create **website location** for evaluation
- establish and frequently utilize **e-networks** for news/knowledge-sharing
- establish **evaluators' forum**; convene periodic events or workshops where evaluators get together
- connect with **evaluators from other countries**

Build a “Best Practices” Archive

- central institution should build an **accessible PE archive**
- **past, related PEs** can be identified as core research documents for upcoming PEs
- **case studies** of lessons learned from successful or unsuccessful PEs can be compiled
- **other processes** may utilize similar methodologies; information banks may exist; (*e.g., ex-ante costing of policy/program proposals*)

Use Comparators

- web links to international sites (e.g., World Bank) with **global indices**
- **international benchmarks**
- similar **evaluations of other countries**
- potential comparators and related evaluations should be identified at evaluation design stage
- similar programs may exist in different sectors (*e.g., anti-smoking, seatbelt use, and pollution reduction ; all have public education and behavioural change dimensions*)

Fill the Gaps

- reality is that **good baselines/data will not exist** for all programs
- in such cases, evaluations may need to rely on more **limited “evidence”** (e.g., questionnaires/ interviews)
- **alternative measures** (e.g., *from customer surveys*) should be explored
- where **data gaps** exist, PE should recommend how to establish a baseline and what data should be collected so that next evaluation will be more evidence-based
- should be **linked to broader performance measurement** approach (*i.e., a results-based management framework*)

Present to Decision Makers

- too many PE reports sit on the shelf
- at a minimum, **Minister should be aware** of the PE before it commences and be presented with the results
- for major PEs, results may be presented to a **broader political audience** (*e.g., group of affected Ministers, government meeting*)
- format for **presenting results** should be developed; Ministers will rarely read full PE

Implement

- credibility of future PEs depends on whether or not the **conclusions of current PEs** are acted on and implemented
- not reasonable to assume that every recommendation is adopted, but PEs **must demonstrate relevance** and impact
- ultimately, the PE evaluation function itself should be evaluated; **are PEs providing value for money?**
- a positive result will, to a large degree, depend on the evidence; **did the PEs make a tangible difference?**