APPENDIX C
GUIDE TO BENCHMARKING OPERATIONS PERFORMANCE MEASURES

Participant Workshop Minutes
February 27 & 28, 2007

NCHRP 20-7
Guide to Benchmarking Operations Performance Measures
Participant Workshop Minutes
February 27 & 28, 2007
National Academies Building, Washington, DC

Those in attendance included:

Glenn Blackwelder	Utah DOT
Ann Brock	TRB
John Collins	Traffic.com
David Crisler	Nevada RTC
Rich Cunard	TRB
Ray Derr	TRB
Walt Diewald	TRB
George Gilhooley	Florida DOT
John Lindquist	Northern VDOT
Tim Lomax	Texas A&M
Rich Margiotta	CSI
Doug Noble	ITE
Jeff Price	VDOT
Rick Shuman	Inrix
Phil Tarnoff	UMD CATT
Rich Taylor	FHWA
Stan Young	UMD CATT
Mike Zezeski	Maryland SHA

<u>via teleconference:</u>

Mary Ameen	North Jersey Trans. Planning Authority
Ali Eghtedari	Vancouver
Eric Hill	Metro Orlando
Valentin Vulov	Georgia RTA

INTRODUCTION

Phil Tarnoff initiated the workshop with a brief review of the results of the National Transportation Operations Coalition (NTOC) Operations Performance Measurement Initiative whose results were published on June 25, 2005. The NTOC initiative consisted of two meetings involving representatives from various organizations including:

- American Association of State Highway and Transportation Officials (AASHTO)
- International City/County Management Association (ICMA)
- Transportation Research Board (TRB)
- Association of Metropolitan Planning Organizations (AMPO)
- American Public Works Association (APWA)
- Institute of Transportation Engineers (ITE)
- University of Maryland Center for Advanced Transportation Technology.

Using the NTOC results as a starting point, the participants reviewed the proposed measures at a high level to determine if any important areas were omitted. In-depth discussions of each performance measure accounted for the majority of the time and effort. Definitions were refined and likely geographic regions (and their corresponding road authorities) were suggested for pilot testing. The workshop concluded with a discussion of guidelines and objectives of the pilot testing phase of the project.

This workshop marks the first major milestone in the NCHRP 20-7 study. Using input from the workshop, each performance measure will be updated. Prior to pilot testing, guidelines for the collection and processing of data will be developed and disseminated to the steering committee for review and comment. The steering committee will consist of workshop attendees, plus other interested parties who were unable to attend. Pilot testing locations will then be identified and data collection is anticipated between May and July of 2007. A final report summarizing all the effort is anticipated by January 1, 2008.

HIGH-LEVEL PERFORMANCE MEASURES DISCUSSION

Discussion Highlights:

- Air quality and safety issues, such as number of accidents, were discussed as possible additions to the list of performance measures. Although there was no disagreement concerning the importance of such matters, due to limits of resources, the NCHRP 20-7 initiative is limited to mobility metrics only.

- The application of the performance measures can apply at various levels. Finer level metrics support the evaluation of individual facilities, while courser level metrics reflect the quality and operation of the system as a whole. The primary motivation of the study is to develop measures to characterize and compare the operation of transportation networks at the system level. However, whenever possible, this effort attempts to provide measures that scale from the base link or

intersection level, up to corridors and regions, and ultimately to system-wide performance.

- Applications that use these measures may be real-time, such as travel times on signs, or solely historical in nature, such as trend-lining of congestion and delay for planning purposes, or anywhere in-between. Suggestions or guidelines related to real-time versus non-real-time applications and supporting data sources are beyond the scope of the study.

- Although the scope encompasses mobility of any form, the application is limited primarily to vehicle travel on highways. This is due in large part to the sponsorship of NCHRP and the participation of supporting organizations, the majority of which are road authorities. The definitions do not restrict the performance measures to highway use only; however, since pilot testing will rely on participating organizations, guidelines and discussion are primarily limited to highway applications.

- The use of traditional engineering measures of highway operations such as Level or Service (LOS) and Volume/Capacity (V/C) ratio are not used directly as operations performance measures. LOS is used as a qualitative condition for determining unconstrained flow conditions. Although LOS and V/C are used frequently in design and engineering of highways, they are specific to highways, do not scale easily, are difficult to measure or observe directly in actual operations, and, probably most importantly, are difficult to communicate to the traveling public.

INDIVIDUAL PERFORMANCE MEASURES REVIEW

The majority of the workshop consisted of in-depth discussions of the proposed performance measures. These are summarized below. The order presented is the order in which they were discussed (and not the order that they appear in the NTOC report). Included in the summary are any recommended changes or clarification to the definitions and suggestions for likely geographic regions and corresponding road authorities where measures may be pilot-tested. [If unfamiliar with the original set of performance measures from the NTOC study, the reader is advised to have a copy of the NTOC final report available for reference.]

CUSTOMER SATISFACTION

Customer satisfaction is frequently debated as a core performance measure. The workshop consensus was to retain the measure, but to address the following issues.

- Explicitly identify the services for which a rating is desired, such as 511, web services, incident management, etc. Develop the actual survey tool using the appropriate expertise such as a university business school or business consultant.

- The survey should contain the following elements:
    - Frequency of use of service (i.e. number of visits per month for a web site)
    - A satisfaction rating
    - Rating of importance or value of service
- All ratings should be based on a numerical scale.
- Make use of existing electronic feedback when available. Web sites and 511 systems provide inherent feedback concerning frequency of use and number of return customers. Although web-based surveys can be more manageable and can easily be used to direct further questions depending on the response of earlier questions (intermediaries), care must be taken to obtain statistically significant results.
- Include questions to the effect:
    - "What else should we be doing?"
    - "Did the information influence your behavior?"

Suggested pilot locations:

Smaller jurisdictions such as MPOs and local municipalities may be able to include the Customer Satisfaction questions in an existing survey. Since MPOs schedule surveys on a periodic basis, they may be able to accommodate the study timeframe.

Puget Sound, Orlando, Southern Nevada, MAG (AZ)
I-95 Corridor Coalition
States: GDOT and VDOT
Also check with the AASHTO Public information officers and AASHTO SSOM

CONGESTION, both SPATIAL and TEMPORAL

Suggested changes to the definition:

- When averaging congestion across multiple roadways, weight the average based on VMT.
- Retain the definition of congestion as a 30% increase in expected travel time.
- Freeways and arterials should be treated separately. Guidelines for establishing expected travel time should address the inherent differences between freeways and arterials.
- Baseline estimate of expected travel-time (corresponding to an unconstrained speed) is based on the 85^{th} percentile of speeds during off-peak traffic.
- Off-peak traffic is defined as times when traffic flows are at level of service C or better.

[Note that the definition of expected travel time, unconstrained speed, and off-peak periods apply to many of the performance measures as noted. These definitions were a primary discussion topic and represent the consensus of the workshop.]

Suggested pilot locations:

Georgia, Florida, Colorado, Maryland (CHART), Virginia DOT
City/MPOs – Orlando, Nevada RTCS, KC

TRAVEL TIME – LINK and TRAVEL TIME RELIABILITY

Suggested changes to the definition:

- Use the term 'Facility' rather than 'Link'
- Both Travel Time and Travel Time Reliability have components of 'Facility' and 'Trip.' Total trip travel time is multi-modal in nature, encompassing not only alternatives such as rail, but also includes end times such as walking, parking and any mode transfers. The 'trip' component of each metric is acknowledged and included in the definition, but available data can only support facility level measurements as they relate to the highway network. Any trip calculation is limited to the aggregation of any respective highway segment travel times. Trip measurements are kept as a placeholder for when adequate data becomes available to support the calculation.
- Reporting interval guidance:
    - For freeways – from interchange to interchange, 2 to 5 miles|
    - For arterials – 0.5 to 2 miles; between intersections of major arterials and/or freeways
    - Between major bus stops
- Travel time metrics should be collected on the finest level possible and then aggregated up to the appropriate reporting levels. For arterials, this may be intersection-to-intersection travel times.
- Unconstrained travel time is based on 85th percentile speed during off-peak periods and off-peak is defined as LOS C or better (same as CONGESTION).

Suggested pilot locations:
Atlanta MPO, Houston, I-95 Corridor Coalition, FHWA Office of Freight Management.

INCIDENT DURATION

Suggested changes to the definition:

- The definition of the end of the incident should be appended with the phrase, "whichever is earlier."
- The measure should be to the number of lanes closed and the type of incident.
- Measure should be based on median time. Incident duration tends to be log-normal in distribution.
- Some events are excluded because the road authorities have no power or influence in shortening the event's duration or influencing its impact.

Suggested pilot locations:
MSHA/CHART, VDOT, I95 Corridor Coalition

RECURRING AND NON-RECURRING DELAY

Suggested changes to the definition:

- Incorporate definitions of unconstrained speed and off-peak hours as in other measures. Unconstrained speed is 85th percentile speed during off-peak periods. Off-peak periods are times when flow is LOS C or better.
- Measures should be collected at the lowest level and then aggregated. For example, record delay on segments, aggregate to roadways, and then to routes.

Suggested pilot locations:

Orlando	ITE	Anaheim	Vancouver	Houston
MNDOT	LA	MD(MMTIS)	DDOT (non-recurring)	

SPEED

Suggested changes to the definition:

- This is a direct reference to spot speed measures – not total distance traveled divided by travel time.
- There was some discussion on whether 'SPEED' should be included as a performance measure. Spot speed prevails in most road authorities as the most directly measurable metric and it allows for trending. The consensus was to retain spot speed as a core performance measure.

Suggested pilot locations:
Piloting is unnecessary.

THROUGHPUT PERSON

Suggested changes to the definition:

- Provides an inter-modal measure of our ability to move people (and/or goods).
- A good measure to reflect effectiveness of corridors or commuting routes.
- No further guidelines are necessary
- Important in evacuation – previous efforts rely on water usage
- Future extensions of this measure may include other forms such as freight.

Suggested pilot locations:
Houston Metro
MPOs conducting screenline surveys – contact AMPO

NTOC Assoc. NAWG

THROUGHPUT VEHICLE

Suggested changes to the definition:

- Aggregate data into a few basic vehicle classifications.

Suggested pilot locations:

See 'Throughput Person'

GENERAL GUIDELINES FOR PILOTING MEASURES

The guidelines and questionnaire for pilot testing of these measures should include the following items:

- Cost of data collection
- Is the data collection already a part of the program?
- Did any data quality issues arise?
- Could the metric be applied consistency between/across jurisdictions?
- Were any processes / procedures used apart from those in the guidelines?
- Are the methods scalable to different sized areas – i.e. New York to Wichita?