1999 Commuter Assistance Program Evaluation Manual

Prepared for:

Florida Department of Transportation

Prepared by:

Center for Urban Transportation Research
University of South Florida
Tampa, Florida

CHAPTER ONE
INTRODUCTION

The Florida Commuter Assistance Program (CAP) is an important and integral part of the Florida Department of Transportation's (FDOT) program to meet transportation needs in the State of Florida. Specifically, the development of a statewide network of CAP offices was completed to offer travel choices to Florida's commuters. According to the official FDOT procedures the Florida Commuter Assistance Program is described as:

"Coordinated use of existing transportation resources can provide a responsive, low cost, alternative for alleviating urban highway congestion, improving air quality and reducing the need for costly highway improvements. The commuter assistance program focuses on the single occupant commuter trip which is the greatest cause of peak hour highway congestion. A coordinated effort to provide alternatives to these commuters using existing or low cost resources, can be beneficial to the development of public transit statewide, the attainment of the Department's program objectives for meeting the transportation needs of the disadvantaged, and the Department's priority efforts to relieve traffic congestion, improve air quality, and to assure energy conservation."

As part of their efforts to ensure that Florida's transportation needs are addressed, the FDOT has specific program requirements for each FDOT District Office and each CAP office. These requirements include establishing specific and achievable program objectives, a listing of tasks to undertake and key activities to perform, reporting on each projects performance including written reports, and measurable goals and objectives with milestones to determine progress in stated emphasis areas. All of these requirements are intended to provide the Department with a tool to evaluate how well CAP offices are meeting FDOT priority efforts to relieve traffic congestion, improve air quality, and to assure energy conservation.

This manual was developed to assist Florida's Commuter Assistance Programs (CAP) in their efforts to measure and evaluate their performance. As such, this manual will focus on providing the information necessary for a CAP to devise and conduct their own evaluation program. It will also provide guidance on how to report the results of that evaluation so that key CAP funders, elected officials, and the general public can understand and appreciate the efforts of the CAP in addressing traffic congestion, air quality, and mobility concerns.

For the ease of use, this manual has been divided into chapters covering specific areas of evaluation. These are:

Chapter Two focuses on the performance measures that a CAP can use to evaluate program progress and record achievements. Included in this chapter are definitions for FDOT required performance measures, FDOT optional performance measures, and a set of other performance measures that a CAP

could use to measure effectiveness and/or report progress. Also included are tables which can be used by a CAP to report results and to track progress.

Chapter Three examines the different types of evaluation that a CAP office may undertake to measure performance and/or progress. Included are descriptions of techniques such as needs assessments, formative evaluation, summative evaluation, and others. Each is described to help the CAP office determine what evaluation is most appropriate to accomplish evaluation objectives.

Chapter Four discusses the different types of survey methodologies that can be used by a CAP office. These include a variety of data collection methods, such as focus groups and mail surveys, as well as sampling considerations.

Chapter Five serves as an introduction to basic statistics. It is intended to provide a working knowledge of statistical principles that can impact a CAP evaluation. The focus is on such items as confidence intervals, statistical differences, and other important characteristics that can impact the quality and reliability of a CAP evaluation program and its results.

Chapter Six addresses survey planning and budgeting. It provides guidance on times at which evaluation is conducted (i.e. season, frequency), examines externalities that may influence the survey, and budgeting issues that must be considered when designing a survey. The chapter also provides guidance on survey costs.

Chapter Seven deals with how evaluation findings can be communicated to those who need to know. This includes a discussion of who needs to know what and when, how to communicate findings, and how to compare CAP findings with other programs.

As each CAP begins to design its own evaluations, it should keep in mind that everyone who examines the evaluation results will bring different expectations and experiences to the review. For example, an MPO may seek to determine how well the CAP is achieving regional transportation objectives. Funders will seek to ensure that funds are being spent in a cost effective manner. To address each of these different expectations, the CAP must carefully design an evaluation that takes into consideration these viewpoints. This manual will provide guidance on important considerations for a CAP that lead to successful evaluations.

CHAPTER TWO
PERFORMANCE MEASURES

Introduction

This chapter will focus on the performance measures available to Florida Commuter Assistance Program (CAP) offices to determine program progress and/or effectiveness. The performance measures are divided into three sections: required performance measures; optional performance measures; and other performance measures. As the name suggests, required performance measures are those that the Florida Department of Transportation (FDOT) Central Office has mandated that all CAP offices in Florida must track and report on at least an annual basis. These performance measures are specified on pages 8-9 of the Commuter Assistance Program procedures, dated May 5, 1997. District optional performance measures are those that FDOT have determined are appropriate for some of the CAP programs and, at CAP and FDOT District option, can be reported to show progress and/or performance. Other performance measures are those that can help a CAP illustrate the effectiveness of their programs in meeting program or regional objectives.

Section A - Required Performance Measures

The FDOT required performance measures are:

1. Number of commuters requesting assistance
2. Number of commuters switching modes
3. Number of vans in service (where applicable)
4. Number of vehicle trips eliminated
5. Vehicle miles eliminated
6. Employer contacts
7. Parking spots saved/parking needs reduced
8. Commuter costs saved
9. Major accomplishments

The following tables have been developed in the CAP evaluation manual to assist the Commuter Assistance Agencies in Florida track their performance relative to FDOT requirements. The tables are constructed with five supporting columns to help the CAP collect, analyze, and disseminate the results of the performance measures. The first column includes the performance measures that are required by FDOT. The second column is used if benchmarks or actual results are available for each performance measure. These benchmarks/results could be taken from survey responses, from past commuter assistance program

evaluation reports, or from data available from other similar CAP programs. The third column can be used if results have been measured over multiple years, and thus a comparison can be made back to the benchmark. The fourth column lists the source for evaluating achievement of the performance measure (i.e. database survey). The fifth column can be used by the commuter assistance program to select targets to achieve for each of the performance measures. The sixth column can be used by CAP staff to explain contributing factors in setting and/or meeting the selected targets.

A separate table describes actions that the CAP agencies take to achieve program goals, or potential activities that could be incorporated to achieve the goal.

Following each of the tables, a brief description of each performance measure is included along with the method to be used to collect the necessary information. Where appropriate, the formula for calculating the performance measure is included.

Required Performance Measures

Performance Measures	Benchmark	Evaluation	Source	Targets*	Cont. Factors
RP1 Number of commuters requesting assistance			Collected by CAP		
RP2 Number of commuters switching modes			Database Survey		
RP3 Number of vans in service			Collected by CAP		
RP4 Number of vehicle trips eliminated			Database Survey		
RP5 Vehicle miles eliminated			Database Survey		

Potential Actions

RA1.1 Provide info to commuters about commute alternatives

RA1.2 Develop matching system

RA1.3 Contract for and/or provide vans for commuting purposes

RA1.4 Develop marketing program to:

a) Promote carpooling b) Promote vanpooling

c) Promote transit use d) Promote walk/bike

RA1.5 Develop employer outreach program

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Required Performance Measures (continued)					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
RP6 Employer contacts			Collected by CAP		
RP7 Parking spots saved/parking needs reduced			Database Survey (based on veh. Trips reduced)		
RP8 Commuter Costs saved			Database Survey (based on veh. miles eliminated)		
RP9 Major Accomplishments			Collected by CAP		

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level

Definitions of Required Performance Measures

RP1 Number of commuters requesting assistance

This is the number of people that request assistance of some sort including:

- Carpool matchlist
- Vanpool matchlist or formation assistance
- Transit route and/or schedule information
- Telecommuting information
- Bicycle route and/or locker/rack information

The CAP offices would track the number of requests received and may want to track requests by type. The information would be reported as part of quarterly and annual progress reports.

RP2 Number of commuters switching modes

This is the number of people that actually use the information you provide to change from their SOV (Single-occupant vehicle)mode to carpooling, vanpooling, transit use, telecommuting, walking and/or bicycling.

This information can be gathered by doing sample survey of commuters assisted on a monthly basis by either phone or mail. Every month contact a random sample of the commuters assisted the previous month to see how many actually used the information provided. Extrapolate survey results to estimate total.

Another possibility is to use an annual survey that measures commute modes before and after joining the agency database.

It is recommended that actual data (rather than data modeled based on the number of commuters in the database and applying a fixed percentage) be used where available.

RP3 Number of vans in service (where applicable)

This measure represents the actual number of commuter vans on the road and/or the number of vanpoolers. These numbers would be collected and reported by the CAP office.

RP4 Number of vehicle trips eliminated

This performance measure is calculated by using follow-up survey data or actual data. To calculate, complete the following steps (Appendix B is a completed sample survey that was used to develop the example below that is highlighted in bold text--in this case a CAP customer who chose vanpooling):

1. If the answer to Question 8 is not 1, 2, or 3, then the total vehicle trips reduced is zero. Go on to the next survey.

Answer is 2 - continue

2. Calculate the *total trips reduced by carpooling after contacting the agency* by calculating the following:

(Question 11 + Question 15) * ((Question 12 + Question 16) -1) /
(Question 12 + Question 16) * (Question 13 + Question 14) *
2 trips/day * 49 weeks/year

(0 days/week + 0 days/week) * (0 trips/day + 0 trips/day - 1) /
(0 trips/day + 0 trips/day) * (0 months + 0 months = 0 years) *
2 trips/day * 49 weeks/year = 0

Questions 13 and 14 should be converted into years, UP TO 1YEAR MAXIMUM, by dividing days by 245, weeks by 49, and months by 12. Since this is an annual measurement, IN NO CASE should the sum of Questions 13 and 14 be greater than 1.

3. Calculate the *total vehicle trips reduced by vanpooling after contacting the agency* by calculating the following:

(Question 19 + Question 23) * ((Question 20 + Question 24) -1) /
(Question 20 + Question 24) * (Question 21 + Question 22) *
2 trips/day * 49 weeks/year

(5 days/week + 0 days/week) * (8 trips/day + 0 trips/day - 1 trip/day) /
(8 trips/day + 0 trips/day) * (8 months = .67 years) *
2 trips/day * 49 weeks/year =
(35 / 8 days/week * .67 years * 2 trips/day * 49 weeks / year) = 287.3 trips

Questions 21 and 22 should be converted into years, UP TO 1 YEAR MAXIMUM, by dividing days by 245, weeks by 49, and months by 12. Since this is an annual measurement, IN NO CASE should the sum of Questions 21 and 22 be greater than 1.

4. Calculate the *total vehicle trips reduced through transit use after contacting the agency* by calculating the following:

(Question 27 + Question 30) * (Question 28 + Question 29) *
2 trips/day * 49 weeks/year

**(0 days/week + 0 days/week) * (0 months + 0 months) *
2 trips/day * 49 weeks/year = 0 trips**

Questions 28 and 29 should be converted into years, UP TO 1 YEAR MAXIMUM, by dividing days by 245, weeks by 49, and months by 12. Since this is an annual measurement, IN NO CASE should the sum of Questions 28 and 29 be greater than 1.

5. Calculate the *total vehicle trips reduced through increase in other means* by calculating the following:

(Question 34 + Question 37) * (Question 35 + Question 36)

**(0 days/week + 0 days/week) * (0 months + 0 months) *
2 trips/day * 49 weeks/year = 0 trips**

Questions 35 and 36 should be converted into years, UP TO 1 YEAR MAXIMUM, by dividing days by 245, weeks by 49, and months by 12. Since this is an annual measurement, IN NO CASE should the sum of Questions 35 and 36 be greater than 1.

6. Sum the results of Steps 3 through 5 to determine the total number of trips reduced after contact with the agency.

Sum = 287.3 trips

To calculate the trips reduced for the entire database:

7. Calculate:

(Sum of the vehicle trips reduced for all the surveys) * (size of rideshare database / number of

surveys completed with members of the rideshare database).

RP5 Vehicle miles eliminated

This performance measure is calculated by using follow-up survey data. To calculate, complete the following steps (refer to Appendix B for the sample completed survey that was used to develop the example):

1. Determine the vehicle trips reduced for *each survey* as described above. (remember that this should be 0 if the answer to Question 8 is not 1, 2, or 3)

Answer is 2 - continue

2. Multiply the result from Step 1 by Question 2 *for each survey*.

287.3 trips * 10 miles = 2873 miles

To calculate VMT reduced for the entire database:

3. Calculate:
$$\frac{\text{(Sum of the vehicle miles reduced for all the surveys)*(size of rideshare database)}}{\text{number of surveys completed with members of the rideshare database).}}$$

RP6 Employer contacts

Report number of employer contacts by the following categories:

Number contacted by letter/fax
Number contacted by phone
Number contacted in person
Number of follow-up calls or visits

When reporting include the number of employees at each site. These figures will be tracked and collected by the CAP staff.

RP7 Parking spots saved/parking needs reduced

This is a performance measure that is calculated by determining the number of people using alternative modes at each employment site. It can also be calculated by taking the number of vehicle trips reduced from a database survey and dividing by 2 trips per day and 245 working days per year.

RP8 Commuter costs saved

This performance measure is calculated by multiplying vehicle miles eliminated by the average cost per mile (AAA uses $.448 per mile, the federal government and State of Florida use $.29 per mile).

RP9 Major accomplishments

This performance measure is a listing of all major CAP programs and/or initiatives and the accomplishments of these projects/initiatives. These may include:

New Transit Services Initiated/Improved
Educational Program Initiated
Transportation Planning Initiatives
Guaranteed Ride Home Projects Initiated
Other Implementation Activities

This information would be tracked and collected by CAP staff.

Section B - District Optional Performance Measures

The FDOT defined District optional performance measures are:

1. Gasoline saved
2. Emissions reduced
3. Information materials distributed
4. Special events
5. Media/community relations

The following tables have been developed in the CAP evaluation manual to assist the Commuter Assistance Agencies in Florida track their performance relative to FDOT requirements. The tables are constructed with five supporting columns to help the CAP collect, analyze, and disseminate the results of the performance measures. The first column includes the performance measures that are required by FDOT. The second column is used if benchmarks or actual results are available for each performance measure. These benchmarks/results could be taken from survey responses, from past commuter assistance program evaluation reports, or from data available from other similar CAP programs. The third column can be used if results have been measured over multiple years, and thus a comparison can be made back to the benchmark. The fourth column lists the source for evaluating achievement of the performance measure (i.e. database survey). The fifth column can be used by the commuter assistance program to select targets to achieve for each of the performance measures. The sixth column can be used by CAP staff to explain contributing factors in setting and/or meeting the selected targets.

A separate table describes actions that the CAP agencies take to achieve program goals, or potential activities that could be incorporated to achieve the goal.

Following each of the tables, a brief description of each performance measure is included along with the method to be used to collect the necessary information. Where appropriate, the formula for calculating the performance measure is included.

Because some of the required performance measures require the CAP to survey their database, a sample survey has been included as Appendix A. This survey provides the basic framework needed to collect all necessary information. The CAP can use this survey, develop one on their own, or use this one as a basis for a more comprehensive survey instrument. Appendix B provides a sample completed survey to show how one database member may answer the survey questions. For assistance in developing surveys, contact the TDM Clearinghouse at the Center for Urban Transportation Research.

District Optional Performance Measures					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
OP1 Gasoline Saved			Database survey data calculation		
OP2 Emissions Reduction			Database survey data calculation		
OP3 Information Materials distributed			Collected by CAP		
OP4 Special Events			Collected by CAP		
OP5 Media/Community Relations			Collected by CAP		
Potential Actions OA1.1 Promote/develop alternative transportation programs. OA1.2 Develop and conduct a community outreach/promotional campaign.					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of District Optional Performance Measures

OP1 Gasoline saved

This performance measure is calculated by multiplying vehicle miles eliminated by the average miles per gallon figure from USDOT/NHTSA. For 1997, average fuel consumption is 0.04 gallons/mile (i.e., 24.4 MPG).

OP2 Emissions reduction

This performance measure is calculated by multiplying vehicle miles eliminated by the emission factors for the CAP service area. Emission factors are available from EPA Mobil Sources Office and are available for Hydrocarbons (HC) , carbon monoxide (CO), and nitrogen oxide (NO). For 1999, the average passenger car emissions were estimated at::

* 2.15 grams/mile of HC
* 19.1 grams/mile of CO
* 2.3 grams/mile of NO

Grams are converted to pounds by multiplying the results of this calculation by .0022.

OP3 Information materials distributed

This performance measure details the number and type of informational materials distributed by the CAP. Informational materials may include but are not limited to:

Brochures
Information Packets
Posters
Surveys

OP4 Special events

This performance measure reports the number and type of special events conducted by the CAP staff to promote and/or encourage commute alternative use. Special events may include but are not limited to:

Transportation Days
Commuter Fairs
Special Promotions

This information would be collected and tracked by CAP staff.

OP5 Media/community relations

This performance measure tracks CAP staff efforts in informing the media and general public about CAP activities and programs. Categories may include but are not limited to:

Number of PSAs aired
Number of newspaper articles
Number of news stories
Number of magazine articles

This information would be collected and tracked by CAP staff.

Section C - Other Performance Measures

The performance measures in this section have been developed to allow a CAP the flexibility to tailor an evaluation program that closely matches program goals and objectives. They have also been developed to measure CAP effects on markets and groups, like employers and the general public, that directly or indirectly are influenced by CAP efforts. The performance measures can be used to develop a more complete profile of direct and indirect effects of CAP program activities on commuter mode choice. For example, the performance measures in this section can be used to determine if advertising campaigns influenced members of the general public to try carpooling without ever contacting the CAP office for assistance. To assist the CAP in selecting appropriate measures from this section, some of the FDOT required and optional performance measures have been repeated under appropriate goals. This provides the CAP with a mechanism to find some performance measures that can help develop a complete picture of CAP efforts.

The following tables have been developed in the CAP evaluation manual to assist the Commuter Assistance Agencies in Florida track their performance relative to FDOT requirements. The tables are constructed with five supporting columns to help the CAP collect, analyze, and disseminate the results of the performance measures. The first column includes the performance measures that are required by FDOT. The second column is used if benchmarks or actual results are available for each performance measure. These benchmarks/results could be taken from survey responses, from past commuter assistance program evaluation reports, or from data available from other similar CAP programs. The third column can be used if results have been measured over multiple years, and thus a comparison can be made back to the benchmark. The fourth column lists the source for evaluating achievement of the performance measure (i.e. database survey). The fifth column can be used by the commuter assistance program to select targets to achieve for each of the performance measures. The sixth column can be used by CAP staff to explain contributing factors in setting and/or meeting the selected targets.

A separate table describes actions that the CAP agencies take to achieve program goals, or potential activities that could be incorporated to achieve the goal.

Following each of the tables, a brief description of each performance measure is included along with the method to be used to collect the necessary information. Where appropriate, the formula for calculating the performance measure is included.

Goal 1 - Increase public awareness					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P1.1 % awareness of CAP among employers			Business survey		
P1.2 # first presentations made			Collected by CAP		
P1.3 # follow-up presentation made			Collected by CAP		
P1.4 % of employers with TDM programs			Business survey		
P1.5 % aided awareness of CAP or CAP number among commuters			General public survey		
P1.6 # of customer inquiries			Collected by CAP		
P1.7 % awareness of CAP promotional materials			General public survey		

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

.

Goal 1 - Increase public awareness

Potential Actions

A1.1 Develop coordinated, consistent marketing program.

A1.2 Develop employer outreach materials on TDM strategies.

A1.3 Plan and conduct kick-off events with employers.

A1.4 Provide technical assistance in establishing employer programs.

A1.5 Establish employer outreach campaign to appoint Employee Transportation Coordinated (ETCs) to involve employers in mobility programs.

A1.6 Host ETC training program.

Definitions of Performance Measures for Goal One

P1.1 % awareness among employers

A measure taken from a business survey . The survey asks if businesses are aware of the commuter assistance program.

P1.2 Number of first presentations made to employers

This is a measure that examines how many presentations were made about rideshare services to area employers. This measure represents initial presentations to employers who have shown an interest in commuter assistance program services. This data would be collected through quarterly reports and year-end evaluation reports.

P1.3 Number of follow-up presentations made to employers

This is a required measure that examines the number of second, third and fourth presentations made to businesses in the CAP service area. This data would be collected from quarterly reports and evaluation reports submitted.

P1.4 % employers with TDM programs

This performance measure represents those employers who have designated an employee transportation coordinator or offer one of the following: compressed work weeks, work at home options, preferential parking, parking shuttles, guaranteed ride home programs, or bus or pool subsidies. Data for this measure would come from a business survey.

P1.5 % aided awareness of Commuter Assistance or Commuter Assistance Number among commuters

This measure examines commuter awareness of the CAP agency and/or the recognition of the telephone number commuters can call to receive assistance. This measure would be collected from the results of the general public survey.

P1.6 Number of customer inquiries

The number of customers who contacted the commuter assistance program during the review period. This measure would be tracked internally by the CAP.

P1.7 % awareness of CAP promotional materials

This measure examines the general public's awareness of CAP promotional materials including highway signs, TV and radio ads, etc. This measure is collected through the general public survey.

Goal 2 - Increase productivity of roadway system					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P 2.1 % of TIP projects related to TDM			Collected by CAP		
P 2.2 % of TIP budget spent on TDM related projects			Collected by CAP		
P 2.3 % increase in average vehicle occupancy			Surveys: Database Gen Pub		
P 2.4 % reduction in vehicle miles of travel from 100% SOV for: 1. Database members 2. General public			Surveys: Database Gen Pub		
P 2.5 % reduction in vehicle trips from 100% SOV among: 1. Database members 2. General public			Surveys: Database Gen Pub		

Potential Actions

A 2.1 Attend and participate in MPO meetings to provide input and guide CAP Services.

A2.2 Develop long range vision, goals and objectives for CAP that are consistent with area-wide transportation network goals and programs.

A2.3 Target MPO selected corridors and roadways for intensive rideshare marketing programs.

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measures for Goal Two

P2.1 % of TIP projects related to TDM

This measure would be calculated by CAP agencies based upon the number of Transportation Improvement Program (TIP) projects related to TDM in local plans vs. the total number of TIP projects.

P2.2 % of TIP budget spent on TDM related projects

This measure would be calculated by local rideshare agencies based upon the total value of TDM related TIP projects vs. total value of all TIP projects.

P2.3 % increase in average vehicle occupancy

This measure examines the increase in vehicle occupancy from one evaluation period to the next. In the table, the baseline figure will be used to help the commuter assistance program calculate the percent change. The measure would be taken from a general public survey and database survey.

P2.4 % reduction in vehicle miles of travel

This measures the percent difference between actual VMT and VMT that would occur if all commuters used an SOV for work trips. The calculation would be done once for database members and once for the general public. To calculate:

(total trips in alternative mode per week) x (duration of alternative mode use)

quantity multiplied by:

$$\frac{\text{(passengers-1/passengers) x (49 weeks per year) x (miles per trip)}}{\text{(total trips per week) x (49 weeks per year) x (miles per trip)}}$$

P2.5 % reduction in vehicle trips

This performance measure would be calculated by taking the total number of trips taken versus the total number of trips that would have been taken assuming all alternative mode users formerly drove alone. The percent reduction figure is derived from a database member survey and the general public survey. To calculate:

(total trips in alternative mode per week) x (duration of alternative mode use) x

quantity multiplied by

$$\frac{\text{(passengers-1/passengers) x (49 weeks per year)}}{\text{(total trips per week) x (49 weeks per year)}}$$

Goal 3 – Decrease Traffic Congestion					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P3.1 % of work trips using alternative mode among: 1. Database members 2. General public			Surveys: Database Gen Pub		
P3.2 # of peak period vehicles per 100 employees			Surveys: Gen Pub Database		
P3.3 VMT reduced: General public Database members			Surveys: Gen Pub Database		
P3.4 Vehicle trips reduced: General public Database members			Surveys: Gen Pub Database		
P3.5 % employers with compressed work week programs among: 1. All employers 2. Targeted employers			Business Surveys		

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Goal 3 - Decrease Traffic Congestion					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P3.6 % employees working a compressed work week among: 1. All employers 2. Targeted employers P3.7 % employers with flextime programs among: 1. All employers 2. Targeted employers P3.8 % employees working a flexible work schedule among: 1. All employers 2. Targeted employers			Business Surveys Business Surveys Business Surveys		
Potential Actions A3.1 Decrease the number of vehicles at activity centers/along corridors A3.2 Increase the use of alternatives among commuters at activity centers/along target corridors A3.3 Develop information on compressed work weeks and flexible work hour programs. A3.4 Conduct workshop on alternative work hour programs for human resource managers.					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measures for Goal Three

P3.1 % of work trips using alternative mode

This performance measure would be calculated by taking the total number of trips made by alternative modes (carpool, vanpool, transit, walk, and bike) and dividing by the total number of trips. The figure would be calculated for both database members and from surveys of the general public.

P3.2 Number of peak period vehicles per 100 employees

This measure can be calculated by CAP agencies by multiplying the inverse of the average vehicle occupancy at a worksite by 100. This measure should be used wherever the commuter assistance program is conducting an employer-based campaign.

Alternatively, this measure can be calculated by multiplying the inverse of the average vehicle occupancy taken from the general public survey and/or the database survey by 100.

P3.3 VMT reduced

This is a performance measure taken from both a general public survey and database member survey. It is calculated by taking the VMT reduced per commuter and multiplying by the number of commuters. The formula for calculating this measure is given under the Definitions of Required Performance Measures section beginning on Page Seven.

P3.4 Vehicle trips reduced

This is a performance measure taken from both a rideshare database member survey and a general public survey. It is calculated by taking the vehicle trips reduced per commuter (respondent) and multiplying by the number of commuters. The formula for calculating this measure is given under the Definitions of Required Performance Measures section beginning on page 4.

P3.5 % employers with compressed work week programs

The percentage of businesses offering a compressed work week schedule as determined by a business survey. Included would be figures for all surveyed employers and those targeted by the CAP. Importance would be determined by CAP focus. In other words, does the CAP provide technical assistance to specific employers, or simply market the concept.

P3.6 % of employees working a compressed work week schedule

A performance measure taken from a business survey, the figure reported represents the % of employees actually participating in a compressed work week program, as reported by the employer. Included would be figures for all employees and for those specifically targeted by the CAP.

P3.7 % employers with flextime programs

The percentage of businesses offering a flextime schedule as reported in a business survey. Included would be figures for all employers and those targeted by the CAP.

P3.8 % of employees working a flextime schedule

A performance measure from a business survey, the figure reported by employers would represents the % of employees actually participating in a flextime program. Included would be figures for all employees and for those who work at targeted employers.

Goal 4 – Improve air quality					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P4.1 Pounds of carbon monoxide reduced			Database survey		
P4.2 Pounds of ozone pollutants reduced			Database survey		
P4.3 Pounds of nitrogen oxide reduced			Database survey		
P4.4 Pollution reductions by mode					
Carpool			Database survey		
Vanpool			Database survey		
Transit			Database survey		
Non-motorized			Database survey		
Potential Actions A4.1 Form carpools. A4.2 Form vanpools. A4.3 Encourage transit use. A4.4 Encourage non-motorized mode usage.					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance

measure is at or above" target level.

Definitions of Performance Measures for Goal Four

P4.1 Pounds of carbon monoxide reduced

Using the results of the VMT calculation, CO reduced is derived by:
(19 grams per mile) x (miles reduced per commuter) x (# of commuters/454 grams per pound).

P4.2 Pounds of hydrocarbons reduced

Using the results of the VMT calculation, hydrocarbon reductions are derived by:
(2.15 grams per mile) x (miles reduced per commuter) x (# of commuters/454 grams per pound .

P4.3 Pounds of nitrogen oxide reduced

Using the results of the VMT calculation, nitrogen oxide reductions are derived by:
(2.29 grams per mile) x (miles reduced per commuter) x (# of commuters/454 grams per pound)..

P4.4 Pollution reductions by mode

Using the above calculations except that reductions are based on VMT reduced by mode.

Goal 5 – Conserve energy resources					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P5.1 % employers with telecommuting program			Business survey		
P5.2 % targeted employers with telecommuting program			Business survey		
P5.3 % employees in a telecommuting arrangement			Business survey		
P5.4 % employees at targeted companies in a telecommuting arrangement			Business survey		
P5.5 % reduction in vehicle miles of travel among: 1. Database members 2. General public			Surveys: Database Gen Pub		
P5.6 Gallons of gasoline saved by alternate mode users among: 1. Database members 2. General public			Surveys: Database Gen Pub		
Potential Actions A5.1 Develop materials on telecommuting. A5.2 Hold a workshop with companies on telecommuting. A5.3 Promote alternative mode use					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measure for Goal Five

P5.1 % employers with a telecommuting program

Taken from a business survey, the percentage of employers who offer a telecommuting option to its employees.

P5.2 % of targeted employers with a telecommuting program

Taken from a business survey, the percentage of businesses that work directly with the CAP or are located within a CAP-targeted activity center who offer a telecommuting option to some of its employees.

P5.3 % of employees in a telecommuting arrangement

Taken from a business survey, the % of employees who have taken a telecommuting option, as reported by employers.

P5.4 % of employees at targeted companies in a telecommuting arrangement

Taken from a business survey, the % of employees who work at targeted companies who have taken a telecommuting option, as reported by employers.

P5.5 % reduction in vehicle miles of travel

This measures the percent difference between actual VMT and VMT that would occur if all commuters used an SOV for work trips. The calculation is done once for database members and once for the general public.

P5.6 Gallons of gasoline saved by alternate mode users

Derived by taking the VMT reduction calculation and multiplying by the average miles per gallon figure for passenger vehicles as reported by USDOT/NHTSA (1997 figure is 24.4 mpg). The figure is derived for database members and for the general public from statistics taken from the database member and general public survey respectively. Gallons of gasoline saved by database members is an Optional Performance Measure as defined by the Florida Department of Transportation in the Commuter Assistance Program Evaluation Manual published by the Center for Urban Transportation Research.

Goal 6 – Improve mobility – Carpools					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P6.1 # persons registered			Collected by CAP		
P6.2 # persons placed in carpools			Database survey		
P6.3 Duration of existing carpools			Database survey		
P6.4 % of trips done by carpool and vanpool			Database survey		
Potential Actions A 6.1 Seek to improve carpool matching program operated by CAP A6.2 Customize brochure on options with survey form. A6.3 Develop "Guide on How to Form a Carpool."					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measures for Goal Six - Carpools

P6.1 Number of persons registered

The total number of persons who are registered in the commuter assistance program database. This number will be developed by the commuter assistance agencies as part of their performance measures.

P6.2 Number of persons placed in carpools

The total number of persons placed into carpools. This would be collected and disseminated as part of the quarterly performance report.

An alternative (and less satisfactory) approach is to calculate the figure from the database survey for both direct and total influence by taking the number of people who switched to carpooling (total) and those who switched where CAP information had some influence (direct).

P6.3 Duration of existing carpools

The average length of time that current poolers have been in their pooling arrangement. This figure is taken from a database members survey.

P6.4 % of trips done by carpool/vanpool

The percentage of all work trips done by carpool and vanpool. This figure is taken from a database member survey and/or a general public survey.

Goal 6 – Improve mobility – Vanpools					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P6.5 # vanpools formed			Collected by CAP		
P6.6 # vanpool riders			Collected by CAP		
P6.7 # vanpool meetings			Collected by CAP		
P6.8 # of vans in service			Collected by CAP		

Potential Actions

A6.4 Meet with representative of transit agencies to strengthen vanpool programs.

A6.5 Make arrangements to obtain vans through purchase or lease (e.g., VPSI).

A6.6 Develop fare structure, arrange for maintenance, prepare marketing materials, and introduce program.

A6.7 Develop "New Start" assistance program to subsidize the cost of 4 empty seats for four Months.

A6.8 Hold presentations with groups of employees who live over 20 miles away from work.

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

.

Definitions of Performance Measures for Goal Six - Vanpools

P6.5 Number of vanpools formed

For this performance measure, the CAP agencies would report the total number of vanpools formed during the review period.

P6.6 Number of vanpool riders

For this performance measure, the CAP agencies would report the total number of vanpoolers as part of their quarterly performance reports.

P6.7 Number of vanpool meetings

For this performance measure, the CAP agencies would report the total number of vanpool meetings held as part of their quarterly performance reports.

P6.8 Number of vans in service

This is an FDOT required performance measure. The CAP agencies would report the number of commuter vans on the road as part of their quarterly performance reports.

Goal 6 – Improve mobility – Bicycle/Pedestrian					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P6.8 % employers with bike racks/lockers P6.9 % employers w/shower/storage P6.10 % commuters using walk or bike to work			Business survey Business survey General public survey		

Potential Actions

A6.9 Develop a program to encourage employers to offer incentives and support for bicycle and pedestrian programs.

A6.10 Meet with area bike coordinators and obtain marketing materials for distribution through employers.

A6.11 Meet with employers to discuss plans.

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measures for Goal Six - Non-motorized

P6.8 % employers with bike racks/lockers

This measure would be taken from a business survey. It represents the percentage of businesses that state that they have bike racks and/or lockers at the worksite.

P6.9 % employers with showers/storage facilities

This measure represents the percentage of employers who offer showers and storage facilities to their employees at the worksite. The figures would be taken from a business survey.

P6.10 % commuters who walk or bicycle to get to work

This measure would be taken from a general public survey and/or database survey. It is the percentage of commuters who use bicycles or walk to work.

Goal 6 – Improve mobility – Transit					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P6.11 % employers purchasing transit passes P6.12 % commuters purchasing transit passes P6.14 % employers with transit subsidy programs P6.15 park n ride lot utilization rates			Collected by CAP Surveys Business survey FDOT/ CAP collected or estimated via GP survey		
Potential Actions A6.12 Increase the number of employers offering transit subsidies to employees. A6.13 Increase the number of employers selling transit passes to employees. A6.14 Encourage/promote the use of Park n Ride lots as a pick-up/drop-off point for pools and/or accessing transit.					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measures for Goal Six - Transit

P6.11 % of employers selling transit passes

This is a question that could be added to rideshare surveys conducted among area businesses. It represents the percentage of local employers that sell discount transit passes to their employees.

P6.12 % of commuters purchasing transit passes

This is a potential performance measure that would be collected in a database member and general public survey. The measure would represent the percentage of survey respondents who purchase transit passes for commuting to work via mass transit vehicles.

P6.13 % of employers with transit subsidy programs

This is a performance measure taken from a survey of businesses. It would represent the percentage of local employers who indicated that they provided financial subsidies to employees traveling on transit vehicles.

P6.14 Park n Ride lot utilization rates

This is another potential performance measure. A site survey could be conducted, or a survey could be conducted to get the information from the general public and/or database members. The result represents either the percentage of parking spaces being used at local Park n Ride facilities or the percentage of the public or members using the facilities.

Goal 7 - Reduce Costs of Auto Ownership					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P7.1 Gasoline costs savings Database General Public P7.2 Auto maintenance savings ($0.13/mile) Database General Public P7.3 Commuter costs saved Database General Public			Surveys: Database Gen. Pub. Surveys: Database Gen. Pub. Surveys: Database Gen. Pub		
Potential Actions A7.1 Develop CAP marketing campaign based on reduced costs A7.2 Implement marketing campaign					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measures for Goal Seven

P7.1 Gasoline costs savings

This performance measure estimates cost savings accrued from not having to purchase gasoline. It is calculated by taking the VMT reduction figure and multiplying by gallons used per mile by the average automobile and the cost per gallon of gasoline. (VMT x gallons/mile x cost/gallon). Average MPG for 1997 was 24.4, and cost per gallon figures are available from local AAA offices. $1/gallon can be assumed.

P7.2 Auto maintenance savings

For this performance measure, the savings are calculated by taking the VMT reduction figure and multiplying by the maintenance costs of an automobile/mile. (VMT x maintenance cost/mile). Maintenance costs are included in the AAA cost per mile figure and generally are about 10-15 cents per mile.

P7.3 Commuter costs saved

This performance measure is calculated by multiplying vehicle miles reduced (or eliminated) by the average cost per mile to operate an automobile (AAA uses $.448 per mile, the federal government and State of Florida use $.29 per mile).

Goal 8 - Improve Economic Viability					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P8.1 Number of parking spaces saved per day P8.2 Cost per trip provided direct influence and total influence			Database survey Database survey		
Potential Actions A8.1 Provide travel choices A 8.2 Provide cost-effective services					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

.

Definitions of Performance Measures for Goal Eight

P8.1 Number of parking spaces saved

This is an adjusted FDOT required performance measure. It is calculated by taking the vehicle trips reduced figure from the database survey divided by 2 trips per day and by 245 working days, but does not take into account influence of the CAP in getting commuters to switch modes.

P8.2 Cost per trip provided (direct and total)

This is a performance measure that is calculated by using the results of the database member survey. The information needed to calculate the cost per trip provided (direct) is:

1. Total carpool and vanpool trips provided per commuter- same measure and calculation method as trips reduced EXCEPT that the size of the pool is not taken into account.

2. Database size.

3. Influence rate per trip for carpool and vanpool- the number of poolers that say their mode choice was influenced by commuter assistance, weighted by the number of trips taken.

4. Annual budget- the budget of the commuter assistance program.

To calculate:

$$\frac{\text{annual budget}}{\text{(total carpool and vanpool trips provided per commuter) x (database size) x (influence rate)}}$$

Calculating the cost per trip provided (total) assumes that all database members that are in a pooling arrangement were, in some way, influenced by the commuter assistance program. The information needed to calculate the cost per trip provided (total) is:

1. Total carpool and vanpool trips provided per commuter- same measure as trips reduced EXCEPT that the size of the pool is not taken into account.

2. Database size.

3. Annual budget- the budget of the Commuter Assistance Program.

To calculate:

$$\frac{\text{annual budget}}{}$$

(total carpool and vanpool trips provided per commuter) x (database size)

Goal 9 - Increase Customer Inquiry					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P9.1 number of customer inquiries			Collected by CAP		
P9.2 number of applications processed			Collected by CAP		
P9.3 % of employers wanting assistance from CAP			Business survey		
Potential Actions A9.1 Develop marketing campaign aimed at reducing costs/congestion					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

.

Definitions of Performance Measures of Goal Nine

P9.1 Number of customer inquiries

The number of customers who contacted the commuter assistance program during the review period. This measure is tracked internally by the CAP agencies.

P9.2 Number of applications processed

This is a performance measure that represents the total number of applications received and processed by the CAP agencies during the review period.

P9.3 % of employers wanting assistance from Commuter Assistance

This is a performance measure taken from a business survey. It represents the percent of businesses responding that stated they would like to be contacted by a CAP agency about establishing an employer TDM program.

Goal 10 - Promote Trial Use					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P10.1 % ever tried alternate mode			Surveys: Database Gen Pub		
P10.2 % of general public trying alternate mode based on advertising			General public survey		
P10.3 % of database trying alternative mode based on CAP info			Database survey		
P10.4 % of general public attempting to contact CAP			General public survey		
Potential Actions A10.1 Develop marketing campaign to encourage use of alternative modes A10.2 Provide rideshare information on request to local residents					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measures for Goal Ten

P10.1 % ever tried alternate mode

This performance measure would be taken from both a general public survey and a database member survey. It represents the percentage of respondents that said they tried using a commute alternative at some point in time to commute to and from work.

P10.2 % of general public trying alternate mode based on advertising

This performance measure is taken from the general public survey. It represents the percent of respondents who said that they tried a commute alternative after hearing/seeing commuter assistance program advertisements.

P10.3 % of database trying alternative mode based on Commuter Assistance info

This performance measure is taken from a database member survey. It represents the percentage of respondents who stated that they tried a commute alternative after obtaining information from the Commuter Assistance Program.

P10.4 % of general public attempting to contact Commuter Assistance

This performance measure would be taken from a general public survey. It represents the percent of respondents who stated that they had tried to contact the CAP agencies for information.

Goal 11 - Facilitate Arrangement of Pools					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P11.1 # of zip code meetings held			Collected by CAP		
P11.2 % database receiving pooling tips			Database survey		
P11.3 % database receiving GRH info			Database survey		
P11.4 % database receiving matching info			Database survey		
P11.5 % database using matchlist to try and form a pool			Database survey		
P11.6 Satisfaction with CAP among database members			Database survey		
P11.7 % database who would recommend CAP			Database survey		
Potential Actions A11.1 Hold zip code meetings at employment sites. A11.2 Make introductory calls to potential matched poolers.					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

.

Definitions of Performance Measures for Goal Eleven

P11.1 Number of zip code meetings held

This performance measure would be tracked by the CAP. It represents the number of meetings held at employment sites to introduce matched employees residing in the same zip code.

P11.2 % database members receiving pooling tips

This measure would be taken from a database member survey. It represents the percent of respondents who stated they had received pooling tips from the commuter assistance program.

P11.3 % database members receiving GRH info

This measure would be taken from a database member survey. It represents the percent of respondents who stated they received guaranteed ride home program information from the CAP.

P11.4 % database members receiving matching info

This measure would be taken from a database member survey. It represents the percent of respondents who stated they had received matching information from the CAP.

P11.5 % of database using the matchlist to try and form a pool

This measure would be taken from a database member survey. It represents the percent of respondents who reported trying to make contacts with others on their matchlist to try and form a pool.

P11.6 Satisfaction with Commuter Assistance among database members

This is a performance measure taken from a database member survey. It represents the satisfaction database members have with services provided by the CAP agencies. Respondents rate agencies on a 1 to 10 scale.

P11.7 % of database members who would recommend Commuter Assistance to others

This is a performance measure that would be taken from the database member survey. It represents the percentage of database members who would definitely recommend commuter assistance to others.

Goal 12 - Reinforce Use of Commute Alternatives					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P12.1 # of GRH rides provided			Collected by CAP		
P12.2 # of registered users in GRH			Collected by CAP		
P12.3 % of database provided with GRH info			Database survey		
P12.4 % of database members receiving follow-up contacts			Database survey		
P12.5 % of employers providing incentives			Business survey		
P12.6 % employers providing GRH			Business survey		
P12.7 % of employers w/ETCs			Business survey		
P12.8 % 12 mo.+ database members using commute alternative			Database survey		
Potential Actions A12.1 Provide GRH program. A12.2 Develop follow-up system.					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

.

Definitions of Performance Measures for Goal Twelve

P12.1 Number of GRH rides provided

This is a performance measure that would be tracked by the CAP agencies. It represents the total number of guaranteed ride home rides provided during the review period.

P12.2 Number registered for GRH

This is a performance measure that would be collected and tracked by the CAP agencies. It represents the total number of persons that have registered for the guaranteed ride home program.

P12.3 % of database provided with GRH info

This measure would be taken from a database survey. It represents the percent of respondents from the entire database that stated they had received guaranteed ride home program information.

P12.4 % of database members receiving follow-up contacts

This measure would be taken from a database member survey. It represents the percent of respondents who reported that they had been contacted by the commuter assistance program as a follow-up to materials that had been sent by commuter assistance

P12.5 % of employers providing incentives

This performance measure would be taken from a business survey. It represents the percent of employers responding that they offered financial subsidies to employees who regularly used the transit system to commute to work.

P12.6 % of employers providing GRH

This is a performance measure taken from a business survey. It represents the percent of employers who reported offering their own guaranteed ride home program to their employees.

P12.7 % of employers w/ETCs

This is a performance measure taken from a business survey. It represents the percent of employers who reported designating their own employee transportation coordinator to assist their employees in finding commute alternatives.

P12.8 % 12 mo.+ database members using commute alternative

This is a performance measure taken from a database member survey. The measure represents the percent of database members whose entry date in the database is greater than 12 months and who report that they are still using a commute alternative.

Goal 13 - Develop CAP Constituency					
Performance Measures	**Benchmark**	**Evaluation**	**Source**	**Targets***	**Cont. Factors**
P13.1 # of complaints			Collected by CAP		
P13.2 Complaints resolved			Collected by CAP		
P13.3 # of testimonials received			Collected by CAP		
P13.4 Employer effectiveness rating of **CAP**			Business survey		
P13.5 **CAP** database satisfaction rating			Database survey		
P13.6 % of database members who would recommend **CAP** to others			Database survey		
Potential Actions A13.1 Develop system to track and resolve complaints. A13.2 Develop system to obtain CAP service users' testimonials.					

* Where performance measures involve surveys or other inferential statistics, targets should be set in the form of "statistically significant increase from prior year" or "95% probability that performance measure is at or above" target level.

Definitions of Performance Measure for Goal Thirteen

P13.1 Number of complaints

This is a potential performance measure for the CAP agencies. The CAP agencies would collect the number of complaints they received in regards to their services.

P13.2 Complaints resolved

This is a performance measure that would be collected and tracked by the CAP agencies. The measure would count the number of complaints resolved by the commuter assistance program to the customer's satisfaction.

P13.3 Number of testimonials received

This is a performance that would be collected by the CAP agencies and would represent the number of testimonials and written recommendations made on behalf of the commuter assistance program.

P13.4 Employer effectiveness rating of commuter assistance

This is a performance measure taken from a business survey. It represents the rating given by employers on the effectiveness of services provided by the CAP agencies. The rating scale is from 1 to 10.

P13.5 Satisfaction with the commuter assistance program among database members

This is a performance measure taken from a database member survey. It represents the satisfaction rating given by respondents on the services provided by the CAP agencies. Respondents would be asked to rate the agencies on a scale of 1 to 10.

P13.6 % of database members who would recommend commuter assistance to others

This is a performance measure taken from a database member survey. It represents the percentage of database members who would definitely recommend the commuter assistance program services to others.

Section D-Determining Appropriate Performance Measures

The CAP office should meet with their local FDOT District representative to select which performance measures will be used to evaluate the program. At a minimum, all required performance measures must be included. At CAP and/or FDOT option, performance measures taken from the optional performance measures section and from the other performance measures section may be included.

Selecting Performance Measures

When selecting performance measures, the CAP and FDOT District offices should consider:

* What performance measures can be used to monitor progress in achieving stated program goals and objectives?

* What performance measures can be used to improve program performance or customer service?

* What performance measures help highlight program accomplishments?

* What CAP programs are important and are not measured through the required performance measures?

* What new initiatives or programs have been added since the last evaluation that should be measured?

* Does the available evaluation budget allow us to conduct other surveys besides the database survey? (See Chapter Six for budget considerations).

Assistance in selecting appropriate performance measures, and in developing survey questions to collect the data needed to assess performance is available from the TDM Clearinghouse located at the Center for Urban Transportation Research (CUTR) at the University of South Florida.

An example methodology for measuring overall program effectiveness and changes in productivity

One of the challenges in evaluating the performance of TDM programs across programs and over time is the diversity of goals and objectives as well as different emphasis areas.

The evaluation should help CAPs enhance their performance through focus on dual, results-oriented goals:

1. delivery of ever-improving value to customers, resulting in greater use of alternatives to the single occupant vehicle by commuters; and

2. improvement of overall CAP operational performance (e.g., lower cost per person served).

The selection of products and services, performance measures, and organizational structure usually depends upon many factors such as the service area, the CAP's stage of development, and employee capabilities. The CAP, in cooperation with their key stakeholders should select which objectives and performance measures best describe its mission and accomplishments.

A successful evaluation will use procedures that determine one or more of the following: (1) the extent to which the program has achieved its stated objectives (e.g., increases in Average Vehicle Occupancy); (2) the extent to which the accomplishment of the objectives can be attributed to the program (direct and indirect effects) (3) the degree of consistency of program implementation to plan (relationship of planned activities to actual activities), and, (4) the relationship of different tasks to the effectiveness of the program (productivity). The following CAP Productivity Index summarizes the CAP's operational performance.

Once the information is collected on performance, awareness and customer satisfaction, the next challenge is how to summarize these diverse factors to give an overall assessment of the program, track progress, and revise objectives.

Using the attached "Productivity Matrix" for the key performance measures or ratios, one can quantify the total impact of the performance measures. Referring to the attached table, the first shaded line would be the actual results of the CAP. The shaded blocks scattered below reflect nearly the same value. The range of values shown are for illustrative purposes only and should be established for each CAP. Level 0 represents the lowest value recorded for the criterion ratio over a recent period of time, in which normal operating conditions existed; nominally the worst ratio reading that might be expected. Level 3 represents operating results indicative of performance proficiency at the time the rating scale is established. The highest level, Level 10, is a realistic estimate of results that can be attained in the foreseeable future (e.g., 3 years) with essentially the same resources that are now available. This could be the benchmark of the industry's best.

By looking up the corresponding "Performance Score" on a scale of 0 to 10 to the right, the CAP can gauge how well the program is doing on that factor. Each score is noted in the shaded line near the bottom of the table. By assigning weights to each factor, the program can recognize those items thought to contribute most to the CAP’s individual program. The CAP and/or FDOT might determine these

weights.

The total "Performance Indicator" score reflects the combined, weighted score of each factor. Changes in this score from period to period will recognize changes in productivity.

PRODUCTIVITY MATRIX

(Example only)

Criterion	Quantity (# Veh. Trips Reduced)	Quantity (# Vans in Service)	Quality (Customer Satisfaction Rating)	Aware-ness	Number of ETCs	TOTAL
CURRENT VALUE	55,232 per yr	6 vans	82% somewhat to very satisfied	50% heard of CAP	32 ETCs	
10	80,000	20	100%	95%	65	
9	75,000	18	97%	90%	60	
8	70,000	16	94%	85%	55	
7	65,000	14	91%	80%	50	
6	60,000	12	88%	75%	45	
5	55,000	10	85%	70%	40	
4	50,000	8	82%	65%	35	
3	45,000	6	79%	60%	30	
2	40,000	4	76%	55%	25	
1	35,000	2	73%	50%	20	
0	30,000	0	70%	45%	15	
SCORE:	5	3	4	1	3	
Weight	20%	10%	30%	10%	30%	100%
Weighted Score	1	0.3	1.2	0.1	0.9	3.5
Change in Productivity ((Total weighted score/3)-1) * 100% =						1.17

CHAPTER THREE
EVALUATION TYPES

Introduction

In order to conduct an effective evaluation, it is necessary to understand what the evaluation is supposed to accomplish. A useful typology of evaluations has been drawn from *The Evaluator's Handbook* published by the Center for the Study of Evaluation at UCLA.

Types Of Evaluation

Three basic types of evaluation exist:

Needs Assessment
Summative Evaluation
Formative (Process) Evaluation

Each of these evaluations uses different types of evaluation tools, including planning or goal-setting meetings, examination of existing data or performance measures, and market surveys. The implementation of each of these tools is described later in the CAP Evaluation manual.
The three types of evaluations are described in detail below:

Needs Assessment

A Needs Assessment is conducted when the program being evaluated is to attempting to determine its goals and objectives. At some point in the organization's life, preferably close to the beginning, organizational goals and objectives must be set. The market that the organization is going to serve and the needs of that market that will be filled by the organization must be clearly identified. Needs Assessments are also called for when the organization perceives that significant change is taking place in its market, either due to new technologies, new patterns of behavior, or other major changes that impact the organization, the way it does business, or the needs that the organization is attempting to meet.

Needs assessments typically use one or more the following evaluation tools:

- Surveys to profile the market, including:
  a) Quantifiable (usually telephone, mail, or panel) surveys to determine size, needs, and to identify and profile the market segments for targeting
  b) Focus groups to better understand the specific needs being served
- Overview of the organization's current capabilities - if applicable (i.e. if the needs assessment is occurring after the organization exists rather than as an initial step in the development of the organization)
- Identification/flowcharting of the organization's current processes - if applicable

- Strategic Planning sessions with upper management

Summative Evaluation

A "Summative" evaluation is one in which the effectiveness of the organization is examined in relation to its goals and objectives. Has the organization met its goals? Is it worth the money that is being spent on it? How well are organizational processes performing? Many elements are used in these types of evaluations: financial records, records of sales or transactions, (in the case of CAP's, records of matches requested and performed, growth of the matchlist database, etc.), examination of performance measures data, and survey research on the market served - often including customer satisfaction surveys. The intent of a summative evaluation is essentially to grade the performance of an organization.

Summative evaluations typically use one or more the following evaluation tools:

- Surveys of the served market, including:
  a) Quantifiable (usually telephone, mail, or panel) surveys to determine impact of the organization on market's behaviors (use of carpools, etc.) and/or to determine organization's customer's satisfaction levels.
  b) Focus groups to better understand the specific problems customers have with the organization - usually done after a quantifiable study
- Examination of organizational data - (i.e. accounting, marketing, and other performance)

Formative or *Process Evaluation*

A Formative Evaluation differs from a Summative Evaluation in that its purpose is to analyze organizational processes and suggest improvements to those processes to better serve the organization's goals - as opposed to merely grading their current effectiveness. The purpose of these evaluations is not so much to find new directions or objectives for the organization to meet as to fine-tune the method currently used in meeting objectives. If there is reasonable doubt that the processes are even coming close to meeting objectives, a summative evaluation of those processes (with the purpose of determining whether or not to continue the activity) may be called for. If there is reasonable doubt that the goals which the process is designed to meet are appropriate, a needs assessment may be called for.

One purpose of conducting a formative evaluation would be to examine the organization's processes as whole. A second purpose might be to compare how processes are carried out in different parts of the organization, such as at different sites. It is not uncommon to discover that two commuter assistance programs operating under a single umbrella, theoretically with the same set of procedures and guidelines, have entirely different ways of handling their customers.

Formative evaluations typically use one or more the following evaluation tools:

- Surveys of the served market, including:
  a) Quantifiable (usually telephone or mail) surveys to determine customer satisfaction with processes, market behaviors and how processes can be better

designed to mesh with those behaviors.

- Focus groups/Personal interviews to better understand how customers use the organization's product or service and the specific needs being served
- Flowcharting of the organization's current processes
- Interviews with employees who carry out the organizational processes being evaluated

Multi-purpose evaluations

Many evaluations are conducted for multiple purposes, particularly for both summative and formative purposes. For instance, it is quite common for a survey of an organization's customers survey to contain elements that both grade the organization on its current performance (summative evaluation) and that inquire into customer opinions about how service can be improved, either implicitly (through customer grading of various organizational processes - low grades need improvement) or explicitly. This is an acceptable, and in many cases desirable, procedure as long as the elements of the evaluation that are being conducted for summative versus formative purposes are clearly delineated.

Market Research and Surveying

This chapter will provide the reader with a brief background on market research and surveying techniques and practices, and how they can be integrated into effective evaluations. It is intended to familiarize the reader with the concepts, terms, and options available in the field of market and survey research. This chapter is intended to provide the reader with enough knowledge to manage and oversee survey research projects. However, just as a manual on TDM strategies would not in itself provide a reader with the knowledge to form and operate a Commuter Assistance Program, this chapter does not in itself provide the tools and knowledge necessary to conduct research projects entirely on one's own. Such abilities are gained with years of classroom instruction and field experience.

Purposes of Doing Market Research Surveys

Market research surveys are designed to answer questions about the attitudes and behaviors of a specific group of people (a "market"), and to provide quantitative estimates of the prevalence of such behaviors and attitudes in the subject population. Market research can be viewed primarily as a means of reducing uncertainty - going from a "guess based on my own experience" to an informed estimate based on interviewing a representative sample of the market in question. A research project can improve an estimate from "I'm pretty sure that somewhere between 20% and 50% of the population has ever actually tried carpooling" to "There is a 90% chance that somewhere between 25% and 30% of the population has ever actually tried carpooling."

Surveys are a tremendous aid in conducting any of the three types of evaluations discussed earlier: Needs Assessments, Summative Evaluations, and Formative Evaluations. They provide greater understanding of how your customers use your products, what they think about them, and what other

products or services they may want that you may be able to provide to them.
Market research projects generally take one of the following forms, described below:

- *Attribute Testing*, which determines what facets or characteristics of a product or service are more or less appealing to a target market. An example of this type of study could be a study on competing airlines: who has better seating, baggage handling, more courteous service, better on-time performance, better prices, etc.
- *Analysis of users*, which provides demographic/psychographic profiles of a target market, often also comparing those profiles to profiles of a different market or of an overall population. This type of study is often used to direct resources in media selection for advertising/promotional campaigns. An example of this type of study could be a comparison of the demographics of carpoolers versus people who drive alone to work.
- *Satisfaction surveys*, which gauge the level of satisfaction of product or service users, and often are also structured to suggest areas where improvement would be most beneficial. An evaluation of a CAP by its members will generally take this form. For that reason, this type of study will be discussed at length in this manual.
- *Studies of decision-making methods*, which investigate how members of a target market make decisions, including what factors are used to make decisions and their relative importance to a decision. An example would be a study of mode choice.
- *Market sizing and/or forecasting*, which attempts to estimate how many people in a target market make use of a product or service and how much of that product they use. An example of this type of study could be an attempt to estimate how a CAP's activities translate into a reduction of Vehicle Miles Traveled.

Whatever the results or findings from a market research study, there are two things that market research never does.

- Research never makes a decision, it merely provides better information for you to make decisions
- Research never guarantees success, it merely reduces the amount of uncertainty in the information you have.

Attribute Testing

The purpose of attribute testing is usually to determine what types of characteristics a product or service should have, and the relative importance of allocating resources to the development, maintenance, or improvement of those characteristics. Respondents are typically asked to rank, rate, or otherwise compare various attributes as to their importance, desirability, value, and so forth. If a rating is used, it is often done on a numerical scale such as 1-5, 1-7, 1-10, etc.

Other types of studies attempt to determine how a product is perceived in terms of its attributes. One such approach, called Multi-Dimensional Scaling, has respondents rate competing products or services in terms of their similarity and then uses mathematical modeling to help identify what

attributes of the products respondents are using to make their comparisons. For example, a survey might have a respondent rate mode choices in terms of their similarity (driving alone versus carpooling versus biking, etc.)
Other techniques have respondents rate each product on a series of attributes and create graphical comparisons of the products based on those ratings. For example, a survey might have respondents rate carpooling, riding the bus, etc., on convenience, cost, efficiency, and so forth.

These types of analyses are often useful in identifying and understand how consumers or potential consumers view competing alternatives, and how perceptions might need to be changed in order to create greater acceptance of a particular alternative.

Analysis of Users

This is a classical type of analysis that usually involves asking respondents about their habits, attitudes, and demographic characteristics (age, income, education, and so forth) and then creates profiles of different groups. Often this is done to identify what types of people are most likely to use a product. This then allows the researcher to try to market the product more actively to those types of people on the basis that it is more attractive to them, or conversely they may try to reposition the product and target it to the types of people who are <u>not</u> using it. It all depends on the purpose of the research and the objectives of the organization that is marketing the product.

Customer Satisfaction Studies

As market growth began to level off and competition for the consumer dollar increased to fierce levels in the latter part of the 1980's and early '90's, customer satisfaction studies grew rapidly in popularity, acceptance, and use. Companies focused more efforts on retaining existing customers as it was discovered that retention was nearly always more efficient and profitable than market expansion and stealing market share from competitors.

Satisfaction studies take on a variety of forms. One of the most common is to measure overall satisfaction with a product or service and also to measure satisfaction with a number of the product's components. For a consumer product such as toothpaste, this might include satisfaction with taste, cleaning ability, cavity prevention, and so forth. For a service, components might include reliability, courtesy of employees, timeliness, and value. Other types of studies measure satisfaction with a large number of different services provided by organizations.

In some cases, statistical models are built that determine the relationship of attribute ratings to overall satisfaction. This can show either what the most important determinants of satisfaction are, or alternatively what elements are most important in explaining the difference between satisfied and unsatisfied customers. The differences here are subtle but extremely important. For instance, in the case of airlines, the most important attribute in customer satisfaction may well be safety, but since

airlines are generally safe (look at the number of accidents compared to, say, roadway accidents) perceptions of safety rarely determine whether or not a customer is satisfied. Other characteristics, such as on-time performance, courtesy of employees, and so forth, become more critical.

Studies of Decision-Making Methods

This is an area that has been heavily used in transportation research. A large number of studies have been done to discover the relative importance of various mode choice determinants (or travel characteristics), including in-vehicle and out-of-vehicle time by mode, perceived costs, parking availability, and so forth. One of the most common approaches is called Discrete Choice Analysis, which is used either with existing data on mode choices (such as census data), or with structured surveys that present respondents with hypothetical situations and ask them to choose a mode given the characteristics of each situation. From the mode chosen and the levels of the characteristics (high parking costs, low parking costs, short travel distance, long travel distance, and so on) the importance of each of the characteristics can be estimated.

Market Sizing and/or Forecasting

This is an extremely common application of survey research, used in many consumer goods and service industries. Respondents are asked to estimate how much of a product or service they use or would use. The sample is then weighted to replicate the make-up of the population in question, and average usage rates are calculated. Finally, these average usage rates are applied to the entire population to determine a total market size or market potential. An example of this type of application could be VMT reduced by getting people to carpool. The population would be surveyed as to their intent to carpool, given some incentives and/or activities that the CAP in the area might undertake. The percentage that would carpool is then reweighted to replicate the make-up of the entire population (if necessary - a well-designed random sampling procedure should just about perfectly replicate the population), and the percentage is then applied to the population size and known travel characteristics. From these calculations overall VMT reduced by forming carpools can be estimated.

This type of procedure has some major limitations. The estimation usually requires respondents to predict entire patterns of behavior of long periods of time, (as opposed to merely stating preference for one product or service over another, or committing to one-time "trial" of a product without long-term implications, which is the form most reliable product/service tests take. Sophisticated demand estimation techniques for products such as consumer goods often use either full-scale test markets or laboratory-based "shops" which allow for observation of behavior and a full representation of the entire choice experience. This type of approach is impossible to apply to carpooling estimation. Carpool estimation also has a relatively rare drawback in that carpooling is seen as a public boon and carpooling is considered socially responsible and desirable. Therefore, respondents are likely to respond that they will carpool when polled as part of a public inquiry, even though their actual behavior will often not follow suit.

Nonetheless, surveying is often the only way of producing a reliable estimate of potential commuting behavior changes. The limitations noted above should be considered when estimations and forecasting are undertaken, but it should also be kept in mind that an estimate with limitations is can be a valuable addition to subjective data and prior experience in other, possibly very different, areas.

CHAPTER FOUR
SURVEY METHODOLOGIES

Types Of Surveys

There are a number of different types of surveys, each of which have unique characteristics and limitations. The choice of survey method is dependent on the objectives involved in doing the project and budget available. The main types of surveys are:

- Focus Groups
- Written/Mail surveys
- Telephone Surveys
- Personal interviews
- Panels

A short discussion of each of the approaches follows:

Focus groups

Focus groups are an excellent alternative if only a very general feel of public interest or support for a particular subject is required, and the researcher wishes to determine which issues of great impact to the community will surface. Because of the small sample sizes involved, this process will not allow for a quantitative estimate of public support nor will it determine the relative importance of issues raised or topics discussed. Typically two to four focus groups will be held. Cost will vary from $3,000 to $6,000 per topic, depending on the number of focus groups held, complexity of questions, and other time-related factors.

Reports from focus groups may contain references such as "75 percent of the people in the group were in favor of....." This type of statement is very misleading, since it implies that the percentage can be applied to the general public. Alternatively, statements such as "there was a consensus that…" will be used, which is also very misleading in that it implies that this consensus will be replicated in the market place. It is best to avoid using numerical results and comparisons if at all possible in such reports, and to concentrate on the qualitative aspects of the results - issues raised and discussed, features of products or services that come up during the session, and so on.

Written/mail surveys

Written and mail surveys are usually the lowest cost alternative available for quantitative estimation. The surveys allow for a relatively large amount of data to be collected from each respondent. However, the format of the questions should be kept simple. Difficult, complex survey formats will usually cause frustration in respondents and low response rates, thereby comprising the sample and possibly rending it

unrepresentative of the population.

Also, written surveys are often subject to low response rates, further compromising projectability. Certain techniques (such as obtaining databases of names and addresses and including incentives) can help to improve response rates at higher costs. Finally, written surveys usually take over a month to collect necessary data.

Costs will vary greatly depending on the level of projectability the researcher is attempting to obtain. To provide a single, reliable estimate for an area, a minimum sample size of 250-300 is recommended. In cases where an independent estimate is required for several segments of the population (such as geographic areas, income levels, etc.), required sample sizes can increase greatly. Usually if a "general idea" is required for sub-segments and an accurate estimate for the population as a whole, a sample size of 125-150 per segment is sufficient. The cost for this type of approach can vary from $5,000 to $10,000 and up, depending on sample size and type required.

Telephone surveys

Telephone surveys have the advantage of rapidly providing quantitative estimates. Telephone surveys also tend to have higher response rates than mail surveys, which increase their level of proper representation and project ability.

The major drawback of telephone surveying is the cost involved. Furthermore, the amount of data and complexity of responses that a respondent can provide is limited - hour phone interviews are not recommended. Concepts presented need to be fairly simple and straightforward.

As with mail surveys, costs will vary greatly depending on the level of projectability the researcher is attempting to obtain. To provide a single, reliable estimate for an area, a minimum sample size of 250-300 is recommended. In cases where an independent estimate is required for several segments of the population (such as geographic areas, income levels, etc.), required sample sizes can increase greatly. Usually if a “general idea” is required for sub-segments and an accurate estimate for the population as a whole, a sample size of 125-150 per segment is sufficient.

The cost for this type of approach can vary from $7,500 to $25,000 and up, depending on sample size and type required and length of interview.

Personal interviews

Personal interviews are the best alternative when complicated survey formats are required and detailed information needs to be provided to respondents. This is the only alternative that would have any chance of providing an estimate of transit demand. However, even this approach would suffer from some of the limitations noted above.

Costs for this type of interview tend to be extremely high if a quantitative estimate is required, since the usual purpose of using this type of interview is to present fairly complex information to potential respondents, and to be able to judge the nuances of response. This requires rather skilled (and relatively expensive) interviewers, and also often involves travel expenses. If the only intent of the personal interview is to be able to present information, a mail/phone approach can sometimes be used at lower cost.

Panels

Panels are used when the objective is to track behaviors and changes in behavior over an extended period of time. Panels also provide convenient samples for testing new ideas in product or service development. Classic examples of panel research include the Nielsen rating panels and a national purchase panel run by the NPD group which tracks purchases of a large number of different consumer goods.

Panel research can be very expensive, particularly if the panel approach is used for a single product or service. Usually panels are most useful when a number of different product or service areas are being covered, as in the NPD panel.

A table summarizing each of these approaches follows:

	FOCUS GROUPS	WRITTEN/MAIL SURVEYS	TELEPHONE SURVEYS	PERSONAL INTERVIEWS	PANELS
Description	8-10 people discuss topics of interest to client; Led by professional moderator	Pre-designed survey mailed out to respondents	Pre-designed survey conducted by professional telephone interviews	Survey administered by individual professional interviewer	Group of respondents who report their behavior over time
Applicable uses	Issue generation; In-depth discussion on complex survey results	General surveying of population; Medium-long surveys; Simpler survey formats	General surveying of population; Short-medium length surveys; Moderately complicated surveys	Interviews with key individuals; Long-Very long surveys; complicated survey formats	Longitudinal studies of behavior and choices; generally standardized survey formats
Costs	Low/Moderate	Moderate	Moderate/High	Very High	Very High
Usefulness for Projections/Trend Analysis	Virtually None - not projectable at all	Only if adequate response rates are obtained	Good	Very Good if enough interviews are completed	Fair - Panel members must be representative
Turnaround	Very fast	Slow	Fast/Moderate	Moderate/Slow	Slow
Strengths	Gets at issues beneath the surface; Low cost; Fast turnaround	Large sample sizes can be obtained; Longer surveys possible	Reasonably representative; fairly good turnaround	Allows more flexibility in interview format, in-depth probing	Allows study of long-term changes in behavior
Weaknesses	Very dependent on having a good moderator; No project ability	Low Response rates/ Unrepresentative samples can occur; Slow turnaround	Higher costs; Surveys need to be kept fairly short and simple	Very high costs per completed survey; Slow turnaround	High cost; Slow turnaround; potential bias based on panel membership
Typical single project cost for complete project (Design, Analysis, Report)	$3,000 - $6,000, based on complexity of issues and number of groups	$5,000 - $10,000 and up, based on complexity of survey and number of respondents	$7,500 - $25,000 and up, based on complexity of survey and # of respondents	$15,000 - $75,000 and up, based on complexity of survey and # of respondents	Varies based on length of study and size of panel

Issues In Sampling

Many of the issues involved in proper sampling have been touched on in the above sections. This section will deal with each of the issues in more depth. The question of sample sizes will be briefly introduced and will covered in detail in the statistics section, which directly follows this section.

Certain key elements that must be included in any sampling plan:

- Definition of target population
- Issues in proper representation
  - a) how to ensure proper representation
    - * quotas & screeners
    - * random selection
    - * reweighting
  - b) how to evaluate how well the sample represents the population
- Sampling efficiency
- Sample size
- Sample sources

Definition of target population

This issue is discussed in the section on hypothesis generation in Chapter 6, "Survey Planning and Budgeting." Usually, the hypotheses that are being tested will define the target population, at least in a broad sense. The key is to define the target population in such a way that each respondent provides meaningful information. Even if the hypotheses do make clear the population that will be surveyed, this item should be restated when the sampling plan is being developed, to ensure that there are no misunderstandings or misinterpretations.

Proper representation

Because most surveys are conducted on a sample of the population rather than the full population, it is vital that the sample selected properly represent the population.

Imagine, for instance, if a survey of potential carpoolers were only conducted among households that had three or more cars, rather than among a sample of the population that more closely represented the entire population. It is very probable that this sample would have very low intentions of carpooling, since car availability is a major factor in determining mode choice. This would lead the researchers to draw erroneous conclusions about the prospects of developing carpools among the population.

How does a researcher go about ensuring proper representation and evaluating completed surveys to check for proper representation?

Ensuring proper representation

Ensuring proper representation can be done in several ways. The steps to be taken include:

1. Identify key variables to serve as indicators
2. Include measurements of those variables in the surveys
3. Devise a random selection process
4. In some cases, require that the sample meet *quotas* on indicator variables
5. Weighting results

1. Identifying key variables- The researcher and the research sponsor should identify those variables that will most likely impact attitudes and behaviors being measured. This is done through a combination of historical data sources (if available) and using the expertise of the parties involved to determine the most important variables. Usually one checks on a limited set of variables, say five or six. These can typically include age, income, gender, presence of children, and so forth.

It is important that there be an independent source that measures those variables. Usually, when the entire population of an area is being surveyed, census data serves as a good check on major demographic variables. Breakdowns of census data or tables in the U. S. Statistical Abstracts can also serve as good checks when segments of a population are being surveyed. When the target sample is from an extremely specific database (for instance, a ridesharing database), data must either be culled directly from the database or from historical surveys of that database, if available.

2. Including measurements of the indicator variables- Clearly, if a variable is to be used as an indicator of proper representation, that variable must be included somewhere in the data collection process. Standard demographics are typically part of any surveying effort, since demographics often impact attitudes and behaviors and are therefore extremely useful in extrapolating results gleaned form a survey to the entire population. Any other variables chosen as indicators, such as number of automobiles, type of housing, and so on, should have a specific question in the survey to collect that data item.

3. Devise a random selection process- The most common way of ensuring a representative sampling of any given target population is through a random sampling process. In telephone-based surveys, this is often accomplished through a technique known as random-digit-dialing. Commercial services will obtain a list of all working phone exchanges, devise a sample of random numbers fitting those exchanges, eliminate exchanges having a high incidence of business/government telephone numbers, and then use the resulting list as a basis for the sample.

This type of list will be most effective if it is further randomized by placing the telephone numbers in random order. Because of the relatively large number of unlisted telephone numbers, a random selection process from a published phone book can create bias by eliminating unlisted numbers (which often belong to people with higher incomes) from the sampling universe.

When sampling from databases is involved, there are several possible random selection procedures. Ideally, the sample will be totally random. The process involved in creating a totally random sample involves:

- determining the sample base necessary
- determining the ratio of sample needed to total database size
- using a random number generator to create numbers between 0 and 1, and applying those numbers to each database record, and
- selecting as sample all those whose assigned random number falls below the ratio of sample needed to database size.

A second, less ideal but more commonly used method, is to create an nth-record sample, where the ratio of database size to sample needed is determined, rounded down, and every nth record is selected, where n is equal to the ratio of database size to sample needed. This method is acceptable when the database is not organized with some sort of regular order bias (such as all database requests sorted by day of the week received).

It should be noted that sample base size, that is, the sample that is drawn to meet the needs of the survey, is usually much larger than the actual required sample size. The reason for this is that there are a large number of non-working phone numbers and/or bad addresses in databases, and that a large percentage of people may not respond to the surveys. A ratio of 10 for sample base to desired completed surveys is not uncommon.

4. Using quotas on indicator variables- Another way of essentially forcing a sample to be representative of the population is to set quotas on some or all indicator variables. This is often used in selecting samples for focus groups, and often used on variables such as male/female ratio and minimum age (often 18 or older) for telephone surveys. Using quotas requires that the indicator variables be identified up front in a portion of the survey called a *screener*. For instance, if a survey were to have quotas set on gender, age, income, and presence of children, where a certain distribution in each of those categories was required, those questions would be the first asked in the survey. Interviewing would take place for each category desired until the quota was filled, and then people meeting the filled-quota description would no longer be interviewed.

A modified form of this approach can be used in mail surveys, but only if many more returns are received than need to be used. The quota variables will be checked on the surveys as they are returned, and as each quota is filled, no more surveys fitting in to that quota will be used. Ideally, this would be done by waiting until a pre-set cutoff date was reached, processing *all* of the surveys received

up to that date, and then randomly selecting surveys to be used for each quota. More commonly, however, quotas will be filled in the order in which the surveys are received. It should be noted that this technique is not often used with mail surveys, except to eliminate returns that don't fit the target population at all. Mail surveys more commonly use weighting techniques to adjust for sample returns, as described in the next section.

5. Weighting survey results- Survey results are commonly *weighted* so that indicator variables will match up with independent source data. For instance, if a survey returned has only a 15% distribution of respondents with 3 or more cars, and it is known that the target population has 25% (say, from census data), then the survey results can be mathematically re-weighted to match the 25% figure. When this is done, *all* of the responses from the 3+ car group are re-weighted, not just the indicator variables. All of their opinions and attitudes are made more prominent.

As an analogy, if you are seeking a medical opinion, and you get one from a doctor who got out of medical school last week and one from a doctor who has been in practice for 10 years, you could reasonably consider all of the statements made by the experienced doctor as being more important to your final decision, on the basis of his/her years of experience. The same principle applies in reweighting survey results.

A critical factor in weighting survey results is that you have sufficient sample size within the group you are reweighting, particularly if you are making their opinions more prominent. If you had 5 responses from people with 3 or more cars and were to weight them as importantly as 100 responses from other people, you run a severe risk of having unrepresentative results. Your confidence in the responses given by the group to be re-weighted should be fairly high. The section on sample size, as well as the section on statistics, will explain the concept of confidence in greater detail. As a rule of thumb, it is probably unwise to re-weight responses from a group with less than 75 respondents.

Evaluating surveys for proper representation

Once the data have been collected, you will have a distribution of responses on the indicator variables, such as percent male and female, percent in various income brackets, and so forth. In some cases, you may have an average (or mean) value as a check (such as mean number of vehicles, mean number of people per household, etc.). Typically, however, indicator variables are evaluated in the form of distributions.

Checking the responses for proper representation essentially involves making statistical tests on the distributions. This section will provide a very general outline of what you are looking for when conducting the tests. The mechanics of conducting the tests will be described in the statistics section.

Two types of tests are commonly conducted on distributions. These tests are a variation on the standard t test and a chi-square test.

The first, in which you compare the percentage of people who fall in a certain category in the survey responses to the percentage that fall in that category in the independent sample (such as the census) is a variation on a standard statistical test called a t test. The t test is designed to be used to compare means. However, in this case, each category can be considered as a yes/no response (for example, if 25% have 3 or more cars, we can treat this as the response to the question "do you have 3 or more cars?" where 25% said yes and 75% said no), and can be essentially treated as a numerical response of 1 or 0. The proportion can then be compared either to historical data or census data, treated in the same fashion, through this test. The mechanics of the test are described in the statistics section.

The chi-square test examines the entire distribution of responses simultaneously, as opposed to comparing category-to-category, and gives back a result that indicates whether the distributions are (statistically) significantly different or not. Thus this test could be applied simultaneously to the percentage of people saying they had no cars, 1 car, 2 cars, and 3 or more cars, to determine if the *entire* distribution were different. Alternatively, it can applied in the same manner as described for the t test (as a series of yes/no responses), in which case the chi-square test is equivalent to the variation on the t test. Again, the notion of statistical significance will be dealt with in detail in the statistics section.

Means can also be compared to ensure representativeness, although this is done much more infrequently. The reason that distributions are used more often to check how whether a sample is representative is that data is the checks are usually done on demographics, which are more typically collected in categorical form rather than in exact numbers.

Sampling efficiency

Collecting data from respondents costs money, and the more data is collected, the more money it costs. Another major cost factor is inefficiency in sampling, where, for example, you set up quotas and then contact a large number of people who don't fit in the quotas. It costs time and money just to check whether or not potential respondents fit into quotas. Usually, research dollars are tight, and it is more than worthwhile to do everything possible to ensure that the sample base is as efficient as possible.

Sampling efficiency can be achieved in many ways. Simple examples could be:

- If a sample of working commuters is desired, it would be wise not to send surveys (or make telephone calls) to communities that are largely populated by retirees.
- If a sample of people who live in, say, St. Petersburg, Florida is desired, all phone exchanges known to be wholly in Clearwater (or Seminole, or Largo, etc.) should be eliminated.

Commercial databases sometimes contain demographic data that can be used. For instance, a survey of commuters drawn from a demographic database could be restricted to those aged 18-54, if age data is available on the database.

For efficiency purposes, if the data is not available in advance and a screener must be used, the

screening section should clearly be the first part of the survey, so that non-qualifying respondents won't be interviewed (and thus cost money), only to determine towards the end of the survey that they don't qualify.

Sample sizes

The issue of how many returned surveys are required is fairly complex. Some fairly advanced statistics are involved. The key issue that the research sponsor needs to determine is the level of uncertainty that is acceptable in the results. As mentioned earlier, there is always a chance that the survey will not exactly represent the opinions of the population even if a completely correct random selection procedure is used. This can be demonstrated with the example of the deck of cards, where we could randomly select 20 cards from the deck and had to estimate (from the cards we drew) what percentage of the cards in the deck were black and what percentage were red. It is conceivable that we would randomly select 20 red cards and no black ones.

Survey results are usually presented as a single, specific result, such as "25% of the population has 3 or more cars." To be completely accurate, the result might be presented in the following way:

> There is a 95% chance that between 22% and 28% (25% +/- 3%) of the population has 3 or more cars. There is a 90% chance that between 23% and 27% (25% +/- 2%) of the population has 3 or more cars. There is an 80% chance that....and so on.

There are two elements involved in the uncertainty about survey results - one is a range of results that the "true" result falls in (known as the *confidence interval)*, and the other is the percent chance that the result falls into that range (known as the *confidence level*). Given a certain sample size that is randomly selected from a population, for any given result - either a percentage or an average - a confidence level and confidence interval can be calculated. The level and the interval are interdependent; that is the size of the interval depends on the magnitude of the level. For any given result, there is an interval corresponding to an 80% confidence level, a different (and larger) interval corresponding to a 90% confidence level, a third (and still larger) interval corresponding to a 95% confidence level, and so forth.

One common misconception is that, in order to get a reliable sample, it is necessary to survey a certain percentage of the population. The fact of the matter is that confidence levels and intervals can be calculated completely independently from the size of the total target population. Should you happen to survey a large percentage of a population (say, 10% or more), a factor can be applied that <u>increases</u> the level of confidence. But the basic calculation (presented in the section on statistics) provides a <u>minimum</u> level of confidence (and confidence interval) independent from the size of the total target population.

The notion of confidence intervals and levels also demonstrates why focus groups are not a reliable source of quantitative information such as percentages. Suppose there are 12 people in a focus group, and eight of them happen to agree on something. It is not uncommon for focus groups to report that "a

large majority" or even "two-thirds" of the "market" agrees on something. Application of the confidence interval formula (which really shouldn't be used for such small samples anyway) would show that the true result, at a 95% confidence level, was anywhere between 41% and 95% - which might not indicate a "large majority" or even a majority at all.

What the research sponsor needs to decide, for the key results coming from the survey, is what size of interval at what level are acceptable. Usually, the confidence level is determined first (e.g., "I want to be 90% confident that all the results..."), and then the acceptable interval is determined ("... are within 3 percentage points or less of the true values.") This decision is then evaluated (using statistics to be presented in the statistics section) for a 50% result, and the desired sample size can then be determined. The nature of the confidence interval is that it is at its maximum size when a 50% result occurs.

Sample sources

There are a large number of potential sources to obtain sample addresses or telephone numbers, whose use depends on the objectives of the survey. These include:

- Databases of rideshare club members
- Commercially available databases drawn from magazine subscription lists, sweepstakes entries, telephone directories, etc. These databases can have a surprisingly large number of names matched to addresses and telephone numbers
- Telephone numbers derived from a random-digits process, which is available from a large number of commercial suppliers
- Databases of business addresses and phone numbers are also available from similar sources.

The choice of which database to use depends primarily on:

- the objectives of the project and the hypotheses being tested
- the extent to which the database covers the target population defined by the objectives and hypotheses. Beware of using databases that are convenient and close at hand but may represent a biased sub-sample of your true target population. For instance, a rideshare database clearly does <u>not</u> represent all carpoolers.
- the expected incidence or "hit rate" expected from the database for efficiency purposes, which is important but must not override the cautions noted just above.

Summary

If all of the above steps are taken, including:

- properly defined target population;
- random selection process;
- checking for proper representation, re-weighting if applicable;
- correct sample size drawn; and
- correct source chosen for the sample;

then the survey should produce reliable information. How useful that information is will depend largely on how well the survey instrument is designed to collect that information. This manual will not attempt to instruct the reader on how to write surveys (which is a skill gained through years of practice and experience).

CHAPTER FIVE
UNDERSTANDING STATISTICS

Introduction

It has been established earlier in this manual that survey research is an effective way to collect information to help evaluate Commuter Assistance Programs. The surveys can produce:

- baseline or benchmark data to which future results will be compared
- results to compare against baseline data
- information about the marketplace which can be used to redirect resources

It should be noted that a survey of a *sample* from a population, rather than a *census* of the population, carries inherent uncertainty. To illustrate the issue, let us return to the example of the deck of cards. Suppose we could randomly select 20 cards from the deck and had to estimate (from the cards we drew) what percentage of the cards in the deck were black and what percentage were red. It is *conceivable*, albeit unlikely, that we would randomly select 20 red cards and no black ones. We would then be forced to conclude (incorrectly, of course) that all of the cards were red.

Statistics

The question that this section will answer is, how much uncertainty arises from a given sampling procedure and how are results analyzed in light of that uncertainty.

Confidence levels and confidence intervals

Two statistical concepts are used to describe the uncertainty arising from a sample:

- Confidence *levels*, which are a measure of the probability that the "true" result lies within a certain range. (The "true" result is the result we would have obtained if we had sampled the entire population rather than just a portion of it)
- Confidence *intervals*, which describe the size of the range mentioned above.

The confidence levels and confidence intervals are dependent on one another. Any given result has a confidence interval associated with a 95% level of confidence, a different (and smaller) interval associated with a 90% level of confidence, another associated with an 80% level of confidence, and so on.

For any given sample, the confidence interval and its associated confidence level can be determined through certain statistical formulas. The formulas may appear daunting at first but they are really quite simple to use. There are several different types of formulas. This section concentrates on the two types used most frequently in survey research:

- those relating to results reported as *proportions* (such as, "25% of the population carpools at

least once per week")

- those relating to results reported as *means* or *averages* (such as, "the average commute distance in the area is 14.6 miles")

While it is not vital for a research sponsor to be able to calculate confidence intervals and perform significance tests, it is a good idea to understand where intervals come from and how tests are performed and what the resulting values mean. This chapter will present the information necessary to make the relevant calculations, and will follow with a table of fairly typical results that should allow the reader to get a general idea of what sort of confidence intervals to expect from data.

Proportions

Given a sample size and a result in the form of a proportion, the confidence interval associated with any given confidence level can be determined.

The first step is to determine the **standard error of the percentage** of the result. In some cases this value has been established, from prior research (such as the census). If the value of the standard error

$$\sqrt{\frac{p(1-p)}{n}}$$

is not known (which is frequently the case), it can be estimated by the following formula:
where: n = size of the sample
p = sample proportion

The standard error is then multiplied by a factor, the value of which is dependent on the confidence level we wish to achieve. Some commonly used values are:

Confidence Level	Factor Value
80%	1.282
90%	1.645
95%	1.960
99%	2.326

These values are valid as long as the associated sample sizes are relatively large (over 30 respondents or thereabouts).

The resulting figure is then added to the survey result to determine the upper limit of the confidence interval, and also subtracted from the survey result to determine the lower limit of the confidence interval.

Using the example mentioned above, suppose a sample of 200 respondents yields the result that 25% (or 40 respondents) carpool at least once per week. Our estimate for how the entire population behaves would then be calculated as follows:

$$\sqrt{\frac{(0.25)(1-0.25)}{200}} = 0.031$$

The confidence interval associated with each confidence level is then calculated by multiplying the standard error by the appropriate factor value:

Confidence Level	Factor Value	Standard Error	Confidence Interval
80%	1.282	0.031	0.040
90%	1.645	0.031	0.051
95%	1.960	0.031	0.061
99%	2.326	0.031	0.072

We can say, therefore that we are 80% confident (or, to be more precise, there is an 80% probability) that the proportion of the population that carpools once per week lies between (0.25-0.04= 0.21 or) 21% and (0.25+0.04= 0.29 or) 29%. This also implies that there is a 20% chance that the proportion of the population that carpools once per week lies between either 0% and 21% or between 29% and 100%. We can furthermore assume that the percent chance of the population's proportion lying in the lower range is equal to the probability of the proportion lying in the upper range, meaning there is a 10% chance of that result being between 0% and 21%, and 10% chance of the result lying between 29% and 100%.

We are 95% confident (or there is a 95% probability) that the population's result lies between (0.25-0.061= 0.189 or) 18.9% and (0.25+0.061=0.311 or) 31.1%, and, as in the example above, we know that there is an equal chance of the result lying above or below those limits, so there is a 2.5% chance that the result is between 0 and 18.9%, and a 2.5% chance that the result is between 31.1% and 100%.

In cases where a significant percentage of the entire target population was surveyed, a factor is applied which increases our confidence in the results. Since the notion of statistical confidence is based on the idea that we might not have surveyed a truly representative sample due to purely random circumstances, it follows that our confidence will increase when we survey a larger percentage of the population, to the point where we are 100% confident if we have in fact surveyed the entire population. This becomes particularly relevant when we sample, for example, rideshare member databases, which might have 800 members and we might survey 250 or so of them.

The factor is calculated by the following formula:

$$factor = \frac{(Total\ Target\ Population\ Size)}{((Total\ Target\ Population\ Size) + Sample\ Size - 1)}$$

The factor is then multiplied by the actual sample size of the survey, and yields what is called the **effective sample size**. This effective sample size, rather than the actual sample size, should be used in all calculations where confidence intervals and analysis of differences require a sample size element.

You will notice from the formula that, unless the sample size is a reasonably large fraction of the target population size, the factor will be virtually equal to 1.

Means

The procedure for determining confidence levels and confidence intervals for results involving a mean value is almost identical to determining levels and intervals for proportions. The only difference is how the standard error is estimated.

Again, the value of the standard error may have been established from prior research. If the value of the standard deviation is not known (which is frequently the case), it can be estimated by the following calculation:

For each observation in the data, calculate:

$$(\text{Result} - \text{Mean of all results})^2$$

which is equivalent to

$$\text{Percentage} * (1 - \text{Percentage})$$

This data element, summed across all observations, is known as the **variance** of the sample.

Then continue by taking the square root of the variance. This is the estimate of the **standard deviation** of the population, and is used in cases where a prior value has not been established. This is equivalent to:

$$\sqrt{Percentage * (1 - Percentage)}$$

Next:

$$\frac{Standard\ Deviation}{\sqrt{Sample\ Size}}$$

This is the **standard error** of the mean.

It is instructive to note that the standard deviation is almost exactly equal to the average difference between each response and the mean value.

The standard error is then multiplied by a factor, the value of which is dependent on the confidence level we wish to achieve. Some commonly used values are:

Confidence Level	Factor Value
80%	1.282
90%	1.645
95%	1.960
99%	2.326

NOTE:

This type of calculation does make one major assumption that was not discussed in the section on percentages. The observed value should be approximately normally distributed, which is to say there should be about ½ of the results above the mean and ½ below the mean, and that there are more results close to the mean than there are far from the mean. A curve of the results should be bell-shaped.

If the results do not follow this pattern, for instance if there are a huge mass of results between 0 and the mean and then fewer, more spread out results above the mean, this type of calculation is inappropriate.

Generally, survey results from larger surveys will follow the assumption of normal distribution. However, it is important to check the results to ensure that this is the case. Particularly with smaller surveys (50 or fewer respondents), the assumption may be violated.

The resulting figure is then added to the survey result to determine the upper limit of the confidence interval, and also subtracted from the survey result to determine the lower limit of the confidence interval.

Using the example mentioned above, suppose a sample of 200 respondents yields the result that the average commute distance is 14.6 miles, and the variance turns out to be 256 miles. Our estimate for

$$\sqrt{(256)} = 16$$

the **standard deviation of the population** would then be calculated as follows:

The **standard error** would be:

$$\frac{16}{\sqrt{200}} = 1.13$$

The confidence interval associated with each confidence level is then calculated by multiplying the standard error by the appropriate factor value:

Confidence Level	Factor Value	Standard Error	Confidence Interval
80%	1.282	1.13	1.45
90%	1.645	1.13	1.86
95%	1.960	1.13	2.21
99%	2.57	1.13	2.63

We can say, therefore that we are 80% confident (or, to be more precise, there is an 80% probability) that the true average commute distance of the population lies between (14.6-1.45=) 13.15 miles and (14.6+1.45=) 15.05 miles.

We are 95% confident (or there is a 95% probability) that the population's result lies between (14.6-2.21=) 12.39 miles and (14.6+2.21=) 15.81 miles.

Table of typical confidence interval sizes at 95% confidence level

Below is a table of typical confidence intervals for means and proportions. 95% has been chosen since it is one of the most widely used confidence levels. The proportions that have been chosen are 10%, 25%, and 50%; the means are on 5-point and 10-point scales with fairly typical standard deviations (which, as was mentioned earlier, are pretty much equivalent to the average difference between each response and the overall mean value).

Keep in mind when using this table that the sample size refers to all respondents answering this question, not necessarily the sample size for the entire project. Some surveys will ask questions of only a portion of the respondents (for instance, "how many people are in your carpool" obviously will only be asked of people who do carpool). Keep in mind that this table also assumes a normal (i.e. bell-shaped) distribution, which is particularly prone to be violated when small sample sizes are used.

Sample Size	10% proportion confidence interval	25% proportion confidence interval	50% proportion confidence interval		5-point scale Average diff. response to mean = 0.8	10-point scale Average diff. response to mean = 2.2
50	1.3%	2.7%	3.5%		0.11	0.31
100	0.9%	1.9%	2.5%		0.08	0.22
150	0.7%	1.5%	2.0%		0.07	0.18
200	0.6%	1.3%	1.8%		0.06	0.16
250	0.6%	1.2%	1.6%		0.05	0.14
300	0.5%	1.1%	1.4%		0.05	0.13
500	0.4%	0.8%	1.1%		0.04	0.10
1,000	0.3%	0.6%	0.8%		0.03	0.07
1,500	0.2%	0.5%	0.6%		0.02	0.06

Determination and analysis of differences for significance

The previous section demonstrated that there is uncertainty about any result that comes from a sample. The "true" result of the target population that was sampled from may not be the same as the result that was obtained from the sample. Statistics allows us to know what is the probable range in which that true result falls.

Now suppose this concept is taken one step further. Suppose we survey *two different* populations, or even one population at two different times, and obtain two results. There will be uncertainty about each of these results, as demonstrated in the previous section. Since we're uncertain about the first result, and uncertain about the second result, they sample results could have come out differently even if both populations had the same "true" result.

For example, suppose we sample one population at two different times, and determined the percentage of commuters who carpooled at least once per week. Suppose in the first sampling we obtained a result at a 95% confidence level of 25% +/- 6.1%, and in the second we obtained a result of 28% +/- 6.1%. Even though the samples both yielded different results, the "true" result could have been 26% in both cases; or it could have been 24% in both cases, or 30%.

If we obtain two results from independent samples, how do we know if the "true" results that they represent are different? The answer comes from an extension of the concept of confidence intervals and confidence levels. If it is possible to determine the percent chance that the "true" result lies within a certain range (for example, in the first of the two carpool results we know that there is a 95% chance that the result lies between 18.9% and 31.1%, a 2.5% chance that the result lies between 0 and 18.9%, and a 2.5% chance that the result lies between 31.1% and 100%, and we know the analogous ranges for the second result), then it should be possible to determine what the chance is that *both* results lie within a certain range for any given confidence level. If we can do that, we can determine what our confidence level is that the "true" results represented by the results of the sample are in fact different. That, in a nutshell, is the concept of statistically significant differences. The rest is applying the appropriate formulas.

Significant differences for proportions

It is not particularly important for research sponsors to comprehend the mathematics behind testing for statistically significant differences. An understanding of the discussion above is quite sufficient. However, for the more mathematically-minded readers, the formulas are presented.

Given two proportion results from two independent samples, the procedure to determine whether or not the proportions are statistically significantly different is:

1. Calculate the value of d:

$$d = \frac{((\text{Sample size 1} * \text{Result 1}) + (\text{Sample size 2} * \text{Result 2}))}{(\text{Sample size 1} + \text{Sample size 2})}$$

2. Calculate the value of the following formula:

$$\frac{(\text{Result 1} - \text{Result 2})}{\sqrt{\frac{(\text{Sample size 1} + \text{Sample size 2})}{(\text{Sample size 1} * \text{Sample size 2})}} * \sqrt{d * (1 - d)}}$$

3. Compare this result to the following table:

If the formula value is <u>at least</u>	The confidence level that the results <u>are significantly different is:</u>
1.282	80%

1.645	90%
1.96	95%
2.57	99%

Significant differences for means

The method for testing for significant differences between mean results follows the same general pattern as the test for proportions:

1. Calculate the variance for each of the two sample results:

$$\text{Sum (Result - Mean of all results)}^2$$

for each observation

2. Calculate the value of the following formula:

$$\frac{(Result\ 1 - Result\ 2)}{\sqrt{\frac{Variance\ 1}{Sample\ size\ 1} - \frac{Variance\ 2}{Sample\ size\ 2}}}$$

3. Compare this result to the following table:

If the formula value is at least	The confidence level that the results are significantly different is:
1.282	80%
1.645	90%
1.96	95%
2.57	99%

Statistically significant differences versus meaningful differences

It is easy to get carried away making calculations of statistical significance of differences, and to lose sight of whether or not those differences are meaningful. Particularly confusing is the question, "is that difference significant?" when what the question really means is, "is that difference meaningful?:

The answer may very well be, "The difference is *statistically* significant, but it isn't *meaningful*. For instance, we might discover that left-handed drivers who ride in carpools drink 1.2 cups of coffee each morning, whereas right-handed drivers who ride in carpools drink 2.8 cups of coffee each morning. Given a reasonable sample size and low variance, this might very well constitute a statistically significant difference. However, while Maxwell House might decide this difference is meaningful, it is doubtful that most CAP managers would find any use for it.

While the above example is admittedly a bit flippant, it demonstrates clearly the difference between significant differences and meaningful differences. This leads back to the discussion at the beginning of the section on formulation of hypotheses. The concepts of confidence intervals, confidence levels, and statistically significant differences allow you to design experiments and test hypotheses that you have made about the population. When the confirmation or denial of the hypotheses leads to re-allocation of resources and effort, the survey has performed its function effectively.

CHAPTER SIX
SURVEY PLANNING AND BUDGETING

Introduction

This chapter will focus on decisions the CAP will have to make before conducting an evaluation. Specifically, the focus of the chapter will be on how to plan and fund an evaluation. While this sounds simple enough, many of the considerations discussed below can have a profound impact on survey costs and data reliability.

Survey Timing

Timing can be a key issue in conducting surveys and can have a significant impact on results if not properly controlled for. In the cable television industry, for example, it is important not to conduct customer satisfaction surveys immediately after rate increases are announced. Employee satisfaction studies are usually not conducted immediately after reviews and/or pay increase announcements for similar reasons. Attitudes towards use of commute alternatives can be affected by prevailing weather patterns, such as extreme heat (or in the case of northern areas, extreme cold). Some elements of timing to be considered when planning surveys include:

Seasonality

Seasonality can be a major issue in survey results, particularly in an area like Florida where there is a high influx of seasonal residents with predictable impacts on traffic levels. Studies evaluating the perceived (or actual) level of congestion will be significantly affected by the season in which they are conducted.

It is not always possible to conduct surveys at "ideal" times, nor is it always possible to determine what an "ideal" time may be. The best approach is usually to do as much as possible to ensure that prevailing conditions are similar when a follow-up survey is conducted. For instance, doing an initial "congestion perception" study during low season, implementing some reduction procedures, and then following up during high season would be methodologically poor, and would probably lead to the conclusion that the policies implemented had actually increased rather than decreased congestion.

Frequency

Survey frequency is another issue that must be dealt with. Budget available is usually a major issue in determining potential survey frequency. Budgets seldom allow for tracking surveys to be conducted more than once a year (if that).

In cases where seasonality may be an issue (see above), you may want to consider spreading your

interview process throughout the year rather than doing all of the interviews at once. This allows for calculation of a rolling average once you have conducted enough interviews to get a baseline, and may give you fairly up-to-the-minute insight into any new situations that may affect your customers or whoever else you are surveying. However, this approach generally involves more expense, particularly if you are having your surveys updated every time you conduct them.

Timing evaluation results for planning and budgeting purposes

Evaluation results are typically desired for year-end evaluations and new year planning purposes. In order to effectively integrate the results of the evaluations into the planning process, the survey must be conducted reasonably far in advance of the planning period. Suggested advance times to start planning the surveys are:

Type of Survey	Advance Time to Start
Focus Groups	2 Months
Mail Surveys	4 Months
Written, hand-distributed surveys	2 Months
Telephone Surveys	3 Months
Personal Interviews	6-8 Months
Panels	N/A, since this is generally an ongoing process.

Budgeting

The primary decision made when budgeting for a survey is the determination of sample size. The concept of how sample size affects the precision of results has been discussed previously. The question that a research sponsor must answer is, how much is the extra precision and certainty from the larger sample size worth?

As a rule of thumb, to get a "quick and dirty" estimate for a population, a sample size of at least 150-200 should be considered. This allows for a wide range of uncertainty, but generally gives a fair idea of the population's attitude.

For a good, solid estimate of the tendencies of a population, sample sizes of 400 or respondents should be considered. Often a sample size of 400 or so may be used to establish benchmarks, and then 200 additional interviews are used as follow-ups to gauge whether there has been any change since the initial study was done.

Planning Survey Projects

Probably the single most important step in planning any research project is the initial planning step. The survey must meet that data needs of the evaluation that you are conducting. If the project is poorly planned in the initial stages, there is virtually no chance that it will result in useful data and meaningful, valuable changes in policy and operations.

The most effective way to plan a research project is to take a rigorous, scientifically-based approach. Ideally, this type of project will be approached as if it were a measurement of a natural phenomenon, as in chemistry, biology, or physics. The basis of the research should be the same as in those sciences. Research design should follow the classic process of hypothesis, experiment, and conclusion. Fortunately for researchers, the types of problems encountered don't demand the analytical complexity of problems in the sciences, but they do demand proper planning and design.

There are five essential elements that any research sponsor must have firmly in mind when initially organizing a research project:

- Given the evaluation being conducted, what decisions will be made with the results of the survey? Or alternatively, how will current operations, policies, and resource allocations be changed based on the survey findings?
- Given the decisions that are being made with the research, what is (are) the specific hypothesis (hypotheses) that is (are) being tested by the research?
- What are the pieces of data that need to be determined in order to make the prove or disprove the hypothesis, and in what form should they be measured? Furthermore, since a sampling process is involved, how confident do we need to be of the results? Is it sufficient for the results to be within 5%, 10%, 50%?
- What are the best sources of information? Does data already exist that answers this question? If not, where is the best place to look for it? If surveying is involved, who are the best people to ask questions of and collect data from?
- How much budget is available to conduct the research?

Each of these areas will be discussed in more detail below:

Step 1: Identify decisions to be made

The evaluation selection process should be a key step in identifying the decisions that are to be made. These decisions should be made explicit at the beginning of the project. This step is unfortunately often omitted from the research process. Even if the evaluator has determined that they will conduct a needs assessment, it is easy to get into trouble by setting vague objectives such as "I want to know what my rideshare database members demographics are." This approach often leads to faulty research design. Often the managers assume that the personnel in charge of actually conducting the research have the same perception of the project's goals, only to find out as the data comes back that some elements

were left out or misinterpreted. Or the research sponsor will assume that he or she understands the process so well that the step of specifying the decisions can be skipped, and the sponsor needs only to ask for specific data elements. This is a serious mistake - the sponsor often discovers new data elements that are needed that could easily have been identified if the planned decisions had been made explicit.
The sponsor should *always* ask for information by specifying the decisions to be made, and *never* merely ask for data. A research sponsor doesn't want to "know the demographics" just to know them. They want to evaluate specific portions of or processes within their organization, or perhaps want to determine which specific actions are required to make the program more effective, such as whether new marketing campaigns are needed, if the entire spectrum of the area's population is being served, and if not, which ones are underserved and why and should resources be allocated to target those groups, and so forth. A simple profile of demographics may or may not provide the data necessary to make those decisions. But if the decisions that are going to be made are known in advance of the design of data-collection instruments and procedures, efficient and correct instruments, sampling plans, and analytical tools can be identified and put to use.

This point cannot be re-iterated too many times. A large number of research projects, possibly even a majority, suffer from a lack of pre-planning and identification of decisions to be made, sometimes to the extent that the entire effort ends up being useless or misleading.

It should be noted that in cases where decisions have been made and will not be changed, due to commitments, regulatory requirements, etc., it is wasteful to spend research dollars to show whether the decision is right or wrong. The research should be directed towards decisions that have not been made and will be made more effectively with additional information at hand.

The decisions that will be made based on the survey results should be explicitly identified by the research sponsor. Will resources be re-allocated and if so, how? If the project is evaluative, how will the evaluation be used to improve operations, policies and procedures, and specifically which operations, policies, and/or procedures are being evaluated? All of this information should be laid out on paper as the first step. Following completion of this effort, the next step is to generate the hypotheses to be tested by the research project.

Step 2: Hypothesis generation

Any experiment in any discipline must test a hypothesis. A research project is an experiment like any other; it should test and either confirm or reject a specific hypothesis (or multiple hypotheses). The hypothesis should take the form of a direct statement, as in "Carpoolers have a significantly different set of demographics than people who drive alone", or "75% of all rideshare database members have a high level of satisfaction with the ridematching service, 'high' being defined as 8, 9, or 10 on a 1-10 scale" The research sponsor should identify the decisions to be made by the evaluation (step 1 above). Then the research sponsor and the research project manager should work together on generating the

hypotheses that, when tested, will provide the sponsor with the information needed for the decisions to be made.

The following elements must be present in any sound hypothesis:

- The *measurement* that is being made and tested (such as a percentage, or an average rating)
- The *scale* that the measurement is being made on (for example, the minimum threshold level where a numerical scale is involved, or the actual statements used in categorical scales)
- The *source*, or *target population*, from which the information will be drawn (such as "rideshare database members" or "all commuters" or "residents of the 5-county area").

If, for example, a re-allocation of resources to target groups that are under-represented in a ridesharing database (compared to the service area's population) is the decision under consideration, one might generate the following hypotheses:

1. The demographics of the ridesharing database are significantly different than the commuter population of the area, specifically in terms of: Income, age, race, gender, presence of children under age 6. (The list might be lengthened, or some elements might be dropped. But the hypothesis should be explicit.)
2. Those demographic groups that are under-represented in the database have a certain minimum threshold interest in carpooling. The minimum threshold interest should also be made explicit: e.g., 20% of the commuters in the area who are in these groups say they are "somewhat or very" interested in carpooling at least once per week on a regular basis. Or one might hypothesize that their interest level is *not* significantly different than the interest level of the demographic groups that are over-represented in the database.
3. One might also generate a hypothesis about the media that would be most useful to use to reach this population. However, it is also quite possible that few media are available (perhaps just direct mail and newspapers) within the budgets allowed, so that regardless of what the research finds, the same approach will be taken. As mentioned above, it is a waste of time and money to identify and collect data for a decision that has already been made and cannot be changed.

The hypothesis should be specific, and should be a direct statement that will either be confirmed or denied by the research. Vague statements like, "Rideshare database members are satisfied with the service provided to them" are not useful or effective hypotheses, because they leave open to interpretation exactly what "satisfied" means. Does this refer to *every* database member? Does it refer to an *average* level of satisfaction, and if so, how is "satisfaction" defined? A better statement would be, "75% of all rideshare database members will say that they are very satisfied (or will rate their satisfaction at least an 8 on a 10-point scale, if a numerical scale will be used) with the ridematching service provided to them."

Step 3: Identification of data needed to prove or disprove hypotheses

Identifying Data Needs

Many research sponsors and research project managers begin their evaluation process at this step, and call it "determining what we need to know." Sometimes, this even takes the form of writing survey questions and specifying response patterns (scales, categories, etc.) without first specifying the type of evaluation being done, what processes or parts of the organization are being evaluated, the decisions to be made with the research, the hypotheses being tested, or the data needed to test the hypotheses, thus greatly compounding the potential for error. As we have seen, it is impossible to effectively determine data needs without having explicit hypotheses. And it should be clear that survey questions should definitely not be written before data needs are determined.

When the hypotheses have been generated, identifying the data needed is actually quite straightforward. By reviewing the hypotheses used above as examples, it is clear that respondent demographics and stated intentions or interests will be included on the questionnaire. It is likely that other hypotheses will have been generated in the planning process as well.

When the data needed have been properly identified, it usually also fairly straightforward for a survey research professional to create the actual survey questions and response scales and/or categories to be used. While it is certainly appropriate for a research sponsor (and presumably this sponsor is not an experienced survey research professional) to review and comment on a questionnaire, it is not advisable for a non-professional to formulate the actual questionnaire. Issues of response bias, question order bias, skip pattern complexity, response choice formatting and design, types and formats of data needed for certain statistical tests and modeling procedures, standard response scaling used in particular types of questions, etc., are all important in questionnaire design but are not issues with which most research sponsors are or need to be familiar.

The Importance of Control Groups

One key concept that is often ignored in evaluations of program effectiveness, particularly where there is a question of what the impact of a program has been, is the notion of a *control group*. A control group is a population that is exactly (or as close to exactly as reasonably possible) like the group on which you are measuring the effects of the program, except that it *has not been exposed* to the program. The measured behavior (such as percentage of people carpooling) should be measured both for the experimental group and the control group to determine what the effectiveness of the program has been. Many experiments skip the step of having a control group by assuming that a control group would have experienced no change in behavior, and thus *any* measured change in the experimental group is due to the program.

This approach can lead to very erroneous conclusions. A major decrease in the price of gasoline, for

instance, may *reduce* the number of people carpooling in the population. If the group that was exposed to the program shows a very small increase in carpooling, it may be concluded that the program was ineffective. However, if it was also known that carpooling within a control group actually *dropped* by 15%-20% due to the decrease in gasoline prices, a different conclusion might very well be reached.

Due to cost constraints, it is sometimes impossible to conduct a research project with an appropriate control group. Other data sources, such as census data, may have to serve as a surrogate for data from a true control group. It is extremely important, however, to understand the notion of a control group and how results from the control group may impact conclusions reached from research data.

The Concept of Sampling

Usually, a research project will involve conducting tests on a sample of the population rather than every member of the population. This occurs because few research sponsors can afford to sample every member of a target population. When this happens, statistical uncertainty is created in the results based on whether the sample accurately represents the population. This is not a question of proper sample design procedures. It is a fact of the sampling process.

To illustrate the issue, the deck of cards example is again instructive. Suppose we could randomly select 20 cards from the deck and had to estimate (from the cards we drew) what percentage of the cards in the deck were black and what percentage were red. It is *conceivable*, albeit unlikely, that we would randomly select 20 red cards and no black ones. We would then be forced to conclude (incorrectly, of course) that all of the cards were red.

Statistical procedures exist that identify what the probability is of having made an error in sampling, and how large that error might be. What must be determined *before* an experiment is undertaken that involves sampling is what level of potential error will be tolerated. This is usually based on the importance and economic ramifications of the decision being made with the research results. This issue was discussed at length in the sections of this manual covering sampling and statistics.

Step 4: Identifying information sources

There are a number of possible sources for information. To determine demographics, for example, there is a wealth of free data available from the U. S. Census. This includes the standard population and housing surveys. In addition, the census releases other, more customizable products, such as the Public Use Microdata Samples (PUMS) which allow the user to create customized cross-tabulations of any census long form data from a 1% sample of all census long forms returned.

Many Commuter Assistance Programs have a number of evaluative tools available from their own records. These include match rates, number of vans in service, number of companies contacted, number of commuters in the database, and so forth. Traffic count data, available from local Department

of Transportation Offices, can also be useful in evaluations and analysis.

In many cases, however, there will be a particular hypotheses that simply can't be proved or disproved by publicly available information, particularly when subjective evaluations (such as satisfaction ratings, ratings of agency responsiveness, and so on) are required. When that situation arises, survey research can provide the means for answering many of these questions. It is therefore imperative that the evaluation planner carefully review all available sources before beginning the survey process.

In a survey research project, it is crucial to ask the right questions. That will be accomplished by carefully following the steps outlined above. It is equally important, however, to ask those questions of the right people. Identifying those people is the crucial first step in developing the sampling strategy. Suppose we determine that we want to estimate the interest level in carpooling among commuters who are not currently in our ridesharing database, as shown in some of the examples above. No matter what questions we ask, we aren't going to get good estimates by interviewing retirees. The goal of the sampling plan should be to identify commuters and interview them and only them. Data from other groups, such as retirees or vacationing families, will not provide data that will help to prove or disprove our hypotheses.

The hypothesis or hypotheses should always give an indication of where to draw the sample from. The hypotheses given above specifically mention "carpoolers" and "rideshare database members." As mentioned earlier, a sound hypothesis should always contain the *source*, or target population, from which the information will come. If the hypothesis is properly constructed, determining the correct population should not be difficult.

Actually obtaining responses from people in those groups and verifying that your respondents did belong to those groups may be more of a challenge. If no available sources exist to pre-identify the people you are contacting as belonging to your target population, it may be necessary to include an identification question (often called a *screener*) in your survey instrument. The screener is essentially a question that verifies the identity of the respondent in relation to the target population. Many surveys have quotas for males and females, for example. Often a research sponsor wishes only to obtain survey responses from adults (18 or older, or 21 or older). If, as in the case above, one only wants to collect data from commuters, a question very early in the survey would ask something like, "Do you commute to work at least three times per week?" to verify that the respondent was in fact in the target population.

Even when a database identifies a person as a member of a target population, it is often a good idea to verify the information through use of a screener. Sometimes databases are out of date or have errors in the entry of data. Using a screener can avoid unnecessary expenditure of usually scarce research dollars on unwanted responses.

Step 5: Determining budget available and the best way to use it

There is often very little leeway in how much budget is available to conduct research. Budget constraints are a very important factor in determining research directions. Limitations on expenditures may eliminate the possibility of conducting certain types of research, or may so limit the number of survey responses you can obtain as to make the information gained of little value. Some objectives may have to be recast in the light of budget realities, particularly in terms of the confidence levels the research sponsor is willing to accept from the data. These considerations must be weighed as the sampling and interviewing plan progresses.

Different types of surveys are available at varying levels of cost. To some extent, the surveys meet different types of objectives. Some survey formats are incompatible with certain objectives. For instance, those with limited budgets may be tempted to use focus groups to prove or disprove quantitative hypotheses (such as, "50% or more of commuters favor HOV lanes over toll roads"). Unfortunately, focus groups are not designed to handle quantitative issues.

Summary

By considering each of the above elements, and following steps in the proper sequence, the effectiveness of research dollars available can be maximized. At a minimum, adherence to the standards set forth in this chapter should maximize the agency's return on their investments in research.

CHAPTER SEVEN
COMMUNICATING EVALUATION FINDINGS

Introduction

While a CAP can take every precaution and devise a nearly flawless evaluation methodology, the value is lost if the CAP cannot effectively communicate the results of their efforts. This chapter will focus on ways in which the Commuter Assistance Programs in Florida can communicate evaluation findings to a variety of audiences.

Getting To Know Your Audience

To develop an effective evaluation report, the CAPs must first understand who their audience is, what information will be of interest to them, and when should the information be available to satisfy that audiences' needs.

Who is the audience for a CAP evaluation report and what do they want to know?

Although the audiences for a CAP evaluation report will differ by CAP, a number of groups with interest in the CAP can be identified. These include the following:

- Funders
- CAP Staff
- CAP Program Directors
- Board of Directors
- Media
- Service Providers
- Politicians
- Clients
- Community Groups
- Other interested parties

Each of these audiences has specific needs from an evaluation. It is up to the CAP to identify what those needs are and to ensure that the information of interest is provided in the evaluation report. Each of these audiences is discussed below.

Funders-An important audience for CAP Evaluation reports. This group will want to ensure that the money provided is being used wisely to achieve identified goals. Prior to beginning an evaluation, the CAP should contact its funders to determine what specific expectations of the CAP program are, and develop an evaluation that answers those questions.

CAP Staff-This is an important audience for CAP evaluation reports because this group is the one that will be most affected by the results. CAP staff can use the evaluation to streamline efforts, to clarify the customer service focus, and to correlate efforts with the achievement of CAP mission and goals.

CAP Program Director-The evaluation should help the director determine if current focus and efforts are achieving desired results. An effective evaluation will help the director refine efforts and target new actions that can help achieve stated goals.

Board of Directors-The evaluation is important to the Board because it helps them determine if their guidance and policy directions are effective in meeting program goals. The evaluation will also help in determining future Board roles.

Media-The media will want two things from an evaluation. They will be interested to see if the CAP is meeting its objectives, and they will want anecdotal information that can be used in developing a story. If anecdotal information is good, the media will develop articles or stories that can be an excellent source of program promotion.

Service Providers-Third party providers, such as taxi companies for guaranteed ride home, can use CAP evaluation results to improve the services provided on behalf of the CAP. Many of these service providers have specific internal customer service and/or satisfaction goals that they want to achieve. The CAP evaluation can help them define their success.

Politicians-The CAP evaluation can help the politician determine if the needs of constituents are being addressed. The evaluation can also serve as an educational/promotional opportunity because it can provide the politician with information about CAP activities and services. Ultimately, the evaluation can serve as a decision-making tool.

Clients-Customers of the CAP are interested in learning about changes in services and how these changes can affect them. They may also be interested in learning how their actions have contributed to the community and/or program success.

Community Groups-Many community groups will be interested in learning what services of the CAP can be beneficial for their success. They may also be looking for ways in which their group and the CAP can work together collectively to achieve common goals. Finally, the community groups may also view the evaluation in the context of comparing their achievements with that of the CAP. This can be especially true if the CAP is a private non-profit that may be competing for funding.

However, when developing an evaluation for a particular set of audiences, the CAP should keep in mind several important considerations. According to Morris, Fitz-Gibbon, and Freeman in "How To Evaluate Evaluation Findings," these considerations are:

- *Different users want different information--even to answer the same question.* A funding agency may accept only valid and reliable test data to prove that a staff training program has been effective, while the personnel participating in the training program would find anecdotal reports and responses from interviews or questionnaires to be the most valid and believable evidence of program effects. Other audiences might require both kinds of information.

- *Some users do not know what they need.* In programs where evaluations are mandated by legal requirements, for example, evaluation clients or program staff may see the assessment simply as a trial to be endured, not necessarily as a process that will lead to useful information and enlightened decisions. If the users are not willing to commit to some criteria for measuring success before the evaluation starts, it is highly unlikely that they will accept or use your final recommendations. Formative evaluators consistently face the task of helping clients define not only program objectives, but also specific evaluation information needs.

- *Some users expect the evaluation to support a specific point of view.* They have already made up their minds about the strengths and weaknesses of the program, and they expect that the evaluation will only confirm their opinions. The results of the evaluation may very well *not* support their preconceptions. So it is vital that the evaluator identify the opinions early on so that he or she can anticipate potential controversies and design reporting procedures which take them into account. Alerting users to your finding discrepancies between their assumptions and the findings as they emerge rather than solely in a final report will make the users more receptive. In fact, an effective evaluation report will contain no surprises, especially with respect to central issues. All of the major questions will have been discussed with program personnel and decision makers from the very beginning, well *before* the final reporting stage. If the evaluation does not bring these issues to light early, the evaluator loses credibility.

- *For some users, the information needs change during the course of the evaluation.* It is not at all uncommon when a formative evaluation is well under way for the users to identify new information they would like to have. Some trainers, for example, might mention that the computer operators in a pilot training program seem to be learning a new data processing system, but the operators have developed a strong dislike for the system. You might change your evaluation plans to include some attitude measures. Although you cannot constantly alter evaluation plans, try to reserve some small portion of your resources to meet requirements for unexpected information that crops up during program implementation.

"How to Communicate Evaluation Findings," by Lynn Lyons Morris, Carol Taylor Fitz-Gibbon, and Marie E. Freeman, Center for the Study of Evaluation, University of California, Los Angeles, CA, pp. 14-15.

As the CAP develops its evaluation, it needs to be aware of these issues and plan accordingly. In most cases, the CAP office will have to decide how to best meet the needs of its primary audience, and develop its evaluation program to meet those needs.

When is the best time to conduct an evaluation?

The simple answer to this question is to say when it will be most useful. The better answer would be to say whenever the evaluation can be used to improve services and the effectiveness of the CAP. In reality, if an evaluation is to be used by all of the potential audiences listed above, then the CAPs would have to continuously evaluate their success. Such an evaluation schedule is impossible, so the CAP should prioritize the most important audiences and complete evaluations to coincide with prioritized needs. Even then, the CAP may need to make some important decisions.

For example, if the purpose of the evaluation is to improve service to justify increased funding, then it stands to reason that the evaluation should be completed to coincide with funding cycles. However, budgets are developed after plans and programs have been determined. This often occurs six months before funding is determined. If the evaluation cannot be used to make improvements to service, or used to determine what services should be offered, then the evaluation may be completed too late to justify increased funding levels that reflect new services.

The following agencies should be contacted in your area to determine when budget and funding decisions are made and when the CAP should be prepared to make its pitch for funds.

- Metropolitan Planning Organization
- Florida Department of Transportation District Office
- Local City and County Governments
- Transit agencies
- Private foundations

With the exception of private foundations, most of the agencies listed above will be on one of two funding cycles, the fiscal year cycle or calendar year cycle. Most fiscal year cycles run July 1-June 30, although federal programs begin a new fiscal year on October 1. As the name implies, calendar year cycles run January 1-December 31.

For private foundations, the exact timing of funding decisions varies greatly and the same foundation may make funding decisions multiple times during the year. For example, the Energy Foundation meets three times a year to review proposals for funding decisions, and requires that materials and proposals be submitted at least eight weeks in advance.

Regardless of who is providing the funds for the CAP, all will probably require an evaluation of efforts. When these evaluation results are due, as well as what will be evaluated and how, should be determined

when the grant is provided. If an evaluation measure is to be tracked internally by the CAP (i.e., number of inquiries about CAP services), the monitoring and/or evaluation should be continuous. This can be especially beneficial if funds are received from FDOT sources who generally require that the CAP include quarterly reports of progress. Again, these requirements will be spelled out when the grant is provided.

Documenting Evaluation Findings

Once evaluations are complete, the CAP must decide how best to convey the results of the evaluation. This is a crucial step that must not be overlooked. A well-designed and carefully managed evaluation can be wasted if the results are not presented in a clear and understandable format. It is also important to remember the potential audiences for the evaluation results and what reporting format will be most useful to meet their needs. The CAP should also be aware that documenting results of evaluations can also be done verbally.

For example, the CAP may be called upon to make a presentation to the County Commissioners on the results of the evaluation. The presentation may be the first exposure the Commission has to the results, and how the results are presented could go a long way in obtaining funding. If the CAP evaluation draws media attention, the results may be broadcast on the radio or television, two media of communication in which written documentation will not be used.

While most CAP offices will commonly be required to disseminate evaluation results in technical reports and/or quarterly progress reports, other forms of communication will typically be used. A list of potential communication media for evaluation results include:

- Technical Report
- Executive Summary
- Brochures
- Press Releases
- Trade Journal Article
- Memorandum
- Public Workshop
- Conference/Seminar Presentation
- Face-To-Face Discussion

Of the audiences for a CAP evaluation report, funders, board members, and CAP staff will have the most interest in a full technical report. Since two of these three audiences have other duties beside CAP oversight, the technical report should be clear and concise, as well as technically credible. A well-written technical report will become a reference manual for this audience.

For politicians, the media, community groups, and clients the preferred written document will be the

executive summary. Even funders and staff will use the executive summary for their own needs. Therefore, the executive summary can be the most important document the CAP will write to disseminate evaluation findings. The summary should be brief, highlight the most important findings of the evaluation, and report the major recommendations of the analysis. Strong support graphics that depict the most important results can be beneficial in the executive summary.

The other communications media listed serve specific audience needs. How the CAP chooses to handle the evaluation findings will dictate which of these media will be used and how they will be used. To strengthen these types of reports, the CAP office should try to determine what evaluation findings are the most important to the audience and focus on preparing a report that best meets that need.

Finally, while the form of communication is important, the CAP must focus its attention on the content of the document. The CAP should:

- Tie together evaluation findings with stated program goals, objectives, and mission of the CAP;
- Compare results to implementation plan and the progress made;
- Demonstrate what effects changes in program offerings have had on service;
- Demonstrate CAP efficiency;
- Examine program strengths and weaknesses;
- Show what problems have arisen, or what trends have changed that may have an impact on results; and,
- Make clear what changes or actions are recommended.

Other important items to consider in the report are:

- Relate information provided to necessary actions;
- Make the report credible;
- Give the audience what it needs, but don't overdo it;
- Present an attractive and readable document;
- Put the most important results first; and,
- Highlight the successes and most important information.

The key for most CAP offices is to look at the evaluation and evaluation report as a powerful tool. If the tool is used effectively it can show the diligence of CAP efforts, the impact the CAP has on meeting community goals and service needs, and the importance of the CAP in solving local and regional problems. A properly planned and well-documented evaluation can be an excellent medium for promoting the CAP and increasing awareness of the community on the important role the CAP plays in Florida municipalities.

APPENDIX A

SAMPLE DATABASE MEMBER SURVEY

CAP Evaluation Rideshare Database Survey

Good evening. My name is _____. I am calling on behalf of the Florida Department of Transportation. This evening we are conducting a short survey on commuting in the *(Insert area name here)* area. We are not attempting to sell you anything, we are only interested in your opinions.

(Ask to speak to person named on sample sheet - repeat intro if necessary)

1. How many days per week do you commute to work?
   _____ (if 0 TERMINATE)

2. And about how far is your commute to work, in miles?_____

3. And at about what time do you usually leave home to go to work?
   1. Before 6 AM
   2. 6 AM – 9 AM
   3. after 9 AM

4. And at about what time do you usually leave work to go home?
   1. Before 3 PM
   2. 3 PM – 7 PM
   3. after 7 PM

5. Have you ever heard of *(Insert name of ridesharing organization here)* ?
   1- Yes 2 - No (Go to END) 9- Don't Know/Refused

6. Have you ever contacted *(Insert name of ridesharing organization here*) for carpool or vanpool information, or not?
   1- Yes 2- No (Go to END) 9- Don't Know/refused

7. Did *(Insert name of ridesharing organization here)* provide you with carpool, vanpool, or transit information or assistance, or not?
   1- Yes 2 - No (Go to END) 9- Don't Know/refused

8. To what extent did the information or assistance provided by *(Insert name of ridesharing organization here)* influence the way you commute to work? Did it:
   1. Have a great deal of influence
   2. Have a moderate influence
   3. Have a slight influence

4. or have no influence at all

9. Did you ever carpool *after* you received the information, or not?
   1- Yes 2 - No (Skip to Q. 17) 9- Don't Know/refused

10. Are you still carpooling to work?
   1- Yes 2 - No (Skip to Q. 14) 9- Don't Know/refused

11. About how many days per week are you carpooling?
   _____ (Enter 0 if question is skipped)

12. About how many people are usually in your carpool, including the driver?
   _____ (Enter 0 if question is skipped)

13. About how long have you been carpooling?
   _____ Days _____ Weeks _____Months _____Years

   [SKIP TO Q. 17]

14. About how long were you in your carpool?
   _____ Days _____ Weeks _____Months _____Years

15. How many days per week were you carpooling?
   _____ (Enter 0 if question is skipped)

16. About how many people were usually in your carpool, including the driver?
   _____ (Enter 0 if question is skipped)

17. Did you ever vanpool to work *after* you received the information, or not?
   1- Yes 2 - No (Skip to Q. 25) 9- Don't Know/refused

18. Are you still vanpooling to work?
   1- Yes 2 - No (Skip to Q.22) 9- Don't Know/refused

19. About how many days per week are you vanpooling?
   _____ (Enter 0 if question is skipped)

20. About how many people are usually in your vanpool, including the driver?
   _____ (Enter 0 if question is skipped)

21. About how long have you been vanpooling?

_____ Days _____ Weeks _____Months _____Years

[SKIP TO Q. 25]

22. About how long were you in your vanpool?
_____ Days _____ Weeks _____Months _____Years

23. How many days per week were you vanpooling?
_____ (Enter 0 if question is skipped)

24. About how many people were usually in your vanpool, including the driver?
_____ (Enter 0 if question is skipped)

25. Did you ever ride the bus to work *after* you received the information, or not?
1- Yes 2 - No (Skip to q. 33) 9- Don't Know/refused (Skip to Q. 31)

26. Are you still riding the bus to work?
1- Yes 2 - No (Skip to Q.29) 9- Don't Know/refused

27. About how many days per week are you riding the bus to work?
_____ (Enter 0 if question is skipped)

28. About how long have you been riding the bus to work?
_____ Days _____ Weeks _____Months _____Years

[SKIP TO Q. 31]

29. About how long were you riding the bus to work?
_____ Days _____ Weeks _____Months _____Years

30. About how many days per week were you riding the bus to work?
_____ (Enter 0 if question is skipped)

31. Is there any other way you used to get to work since you received the information?
1- Yes 2 - No (Go to END) 9- Don't Know/refused

32. And how were you getting to work? (Specify _____)

33. And are you still getting to work by (INSERT ANSWER TO Q. 32)?
1- Yes 2 - No (Skip to Q.34) 9- Don't Know/refused

34. About how many days per week are you (INSERT ANSWER TO Q. 32)?
_____ (Enter 0 if question is skipped)

35. About how long have you been (INSERT ANSWER TO Q. 32)?
_____ Days _____ Weeks _____Months _____Years

[GO TO END]

36. About how long were you getting to work by (INSERT ANSWER TO Q. 32)?
_____ Days _____ Weeks _____Months _____Years

37. About how many days per week were you getting to work by (INSERT ANSWER TO Q. 32)?
_____ (Enter 0 if question is skipped)

END Thank you very much for your cooperation in this survey. Good night.

APPENDIX B

Sample Completed Rideshare Database Survey

CAP Evaluation Rideshare Database Survey

Sample Completed Survey

Good evening. My name is _____. I am calling on behalf of the Florida Department of Transportation. This evening we are conducting a short survey on commuting in the *(Insert area name here)* area. We are not attempting to sell you anything, we are only interested in your opinions.

(Ask to speak to person named on sample sheet - repeat intro if necessary)

1. How many days per week do you commute to work?
   __5__ (if 0 TERMINATE)

2. And about how far is your commute to work, in miles?___10_____

3. And at about what time do you usually leave home to go to work?
   1. Before 6 AM
   2. 6 AM – 9 AM
   3. after 9 AM

4. And at about what time do you usually leave work to go home?
   1. Before 3 PM
   2. 3 PM – 7 PM
   3. after 7 PM

5. Have you ever heard of *(Insert name of ridesharing organization here)* ?
   1- Yes 2 - No (Go to END) 9- Don't Know/Refused

6. Have you ever contacted *(Insert name of ridesharing organization here*) for carpool or vanpool information, or not?
   1- Yes 2- No (Go to END) 9- Don't Know/refused

7. Did *(Insert name of ridesharing organization here)* provide you with carpool, vanpool, or transit information or assistance, or not?
   1- Yes 2 - No (Go to END) 9- Don't Know/refused

8. To what extent did the information or assistance provided by *(Insert name of ridesharing organization here)* influence the way you commute to work? Did it:
   1. Have a great deal of influence
   2. Have a moderate influence
   3. Have a slight influence

4. or have no influence at all

9. Did you ever carpool *after* you received the information, or not?
   1- Yes **2 - No (Skip to Q. 17**) 9- Don't Know/refused

10. Are you still carpooling to work?
    1- Yes 2 - No (Skip to Q. 14) 9- Don't Know/refused

11. About how many days per week are you carpooling?
    __0____ (Enter 0 if question is skipped) (**skipped)**

12. About how many people are usually in your carpool, including the driver?
    __0___ (Enter 0 if question is skipped) **(skipped)**

13. About how long have you been carpooling?
    _____ Days _____ Weeks _____Months _____Years

    [SKIP TO Q. 17]

14. About how long were you in your carpool?
    _____ Days _____ Weeks _____Months _____Years

15. How many days per week were you carpooling?
    __0____ (Enter 0 if question is skipped) **(skipped)**

16. About how many people were usually in your carpool, including the driver?
    __0___ (Enter 0 if question is skipped) **(skipped)**

17. Did you ever vanpool to work *after* you received the information, or not?
    1- Yes 2 - No (Skip to Q. 25) 9- Don't Know/refused

18. Are you still vanpooling to work?
    1- Yes 2 - No (Skip to Q.22) 9- Don't Know/refused

19. About how many days per week are you vanpooling?
    ___5__ (Enter 0 if question is skipped)

20. About how many people are usually in your vanpool, including the driver?
    __8__ (Enter 0 if question is skipped)

21. About how long have you been vanpooling?

_____ Days _____ Weeks ___8___Months _____Years

[SKIP TO Q. 25]

22. About how long were you in your vanpool?
_____ Days _____ Weeks _____Months _____Years

23. How many days per week were you vanpooling?
___0___ (Enter 0 if question is skipped) **(skipped)**

24. About how many people were usually in your vanpool, including the driver?
___0__ (Enter 0 if question is skipped)**(skipped)**

25. Did you ever ride the bus to work *after* you received the information, or not?
1- Yes **2 - No** (Skip to q. 33) 9- Don't Know/refused (Skip to Q. 31)

26. Are you still riding the bus to work?
1- Yes 2 - No (Skip to Q.29) 9- Don't Know/refused

27. About how many days per week are you riding the bus to work?
__0____ (Enter 0 if question is skipped) **(skipped)**

28. About how long have you been riding the bus to work?
_____ Days _____ Weeks _____Months _____Years

[SKIP TO Q. 31]

29. About how long were you riding the bus to work?
_____ Days _____ Weeks _____Months _____Years

30. About how many days per week were you riding the bus to work?
__0____ (Enter 0 if question is skipped) **(skipped)**

31. Is there any other way you used to get to work since you received the information?
1- Yes **2 - No** (Go to END) 9- Don't Know/refused

32. And how were you getting to work? (Specify _____)

33. And are you still getting to work by (INSERT ANSWER TO Q. 32)?
1- Yes 2 - No (Skip to Q.34) 9- Don't Know/refused

34. About how many days per week are you (INSERT ANSWER TO Q. 32)?
__0____ (Enter 0 if question is skipped) **(skipped)**

35. About how long have you been (INSERT ANSWER TO Q. 32)?
_____ Days _____ Weeks _____Months _____Years

[GO TO END]

36. About how long were you getting to work by (INSERT ANSWER TO Q. 32)?
_____ Days _____ Weeks _____Months _____Years

37. About how many days per week were you getting to work by (INSERT ANSWER TO Q. 32)?
__0____ (Enter 0 if question is skipped) **(skipped)**

END Thank you very much for your cooperation in this survey. Good night.

APPENDIX C

Commuter Assistance Program Procedures

Table of Contents

(continued)

Table of Contents (continued)

(continued)

Table of Contents (continued)

(continued)

Table of Contents (continued)