UC Berkeley
Research Reports

Title

Estimating Pedestrian Accident Exposure: Protocol Report

Permalink

https://escholarship.org/uc/item/8j8685jt

Authors

Greene-Roesel, Ryan
Diogenes, Mara Chagas
Ragland, David R

Publication Date

2007-03-01

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Protocol Report

MARCH 2007

Dan Burden

Traffic Safety Center
Setting New Directions in Traffic Safety

The mission of the UC Berkeley Traffic Safety Center is to reduce traffic fatalities and injuries through multi-disciplinary collaboration in education, research, and outreach. Our aim is to strengthen the capability of state, county, and local governments, academic institutions, and local community organizations to enhance traffic safety through research, curriculum and material development, outreach, and training for professionals and students.

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE
Protocol Report

Prepared for Caltrans under
Task Order 6211

Prepared by
RYAN GREENE-ROESEL
MARA CHAGAS DIOGENES
DAVID R. RAGLAND

University of California Traffic Safety Center

University of California
Berkeley, CA 94720
Tel: 510/642-0655
Fax: 510/643-9922

ACKNOWLEDGEMENTS

LIST OF CONTENTS

LIST OF TABLES AND FIGURES

1. PREFACE

1.1. Purpose of the Protocol

Walking is a healthful, environmentally benign form of travel, and is the most basic form of human mobility. Walking trips account for more than 8 percent of all trips taken in California, making walking the second most commonly used mode of travel after the personal automobile (Caltrans, 2002). In addition, many trips made by vehicle or public transit begin and end with walking.

In spite of the importance and benefits of walking, pedestrians suffer a disproportionate share of the harm of traffic incidents in California. As noted above, walking trips make up just 8 percent of all trips in the state, but 17 percent of all traffic fatalities are suffered by pedestrians. In 2004, 694 pedestrians were killed in the state of California and 13,892 were injured (California Highway Patrol, 2004).

To address this problem, significant resources are focused on countermeasures that aim to reduce the risk of pedestrian injury. Because resources are limited, risk analysis is necessary to develop cost-effective countermeasures (Høj and Kröger, 2002).

In the field of pedestrian safety, risk analysis involves assessing factors that contribute to the danger that a pedestrian is struck by a vehicle. These factors may include physical characteristics of the street, such as lack of sidewalks; behavioral issues, such as pedestrian or driver alcohol use; as well as other environmental variables. In order to fully understand how these factors contribute to risk, it is necessary to collect information on pedestrian exposure. Collection of pedestrian exposure information is an essential component of risk analysis.

Pedestrian exposure is a concept that refers to the amount that people are exposed to the risk of being involved in a traffic collision. In principle, pedestrians are exposed to this risk whenever they are walking in the vicinity of automobiles. There are many metrics that can be used to measure pedestrian exposure, but pedestrian volumes are the most frequently used.

Although many state, regional, and local agencies have developed methodologies to collect pedestrian volume data, there is no consensus on which method is best (Schneider et al., 2005; Schweizer, 2005). This is because there is no "one size fits all" method of counting pedestrians. Rather, the choice of strategy depends on a complex range of factors, including the characteristics of the area being studied; the resources available for data collection; and the specific purpose of data collection.

This protocol aims to improve pedestrian data collection in the state of California by providing information and guidance for each decision point in the data collection process. Each chapter represents one of these decision points, and each will guide the user through important considerations relevant to the data collection stage. In addition, each chapter provides a combination of real-world and hypothetical example scenarios to illustrate the issues discussed in the text.

The first chapter, "Pedestrian Exposure," discusses the issue of how to select a definition of pedestrian exposure that is appropriate to the study purposes, resources, and chosen counting method. It also discusses the meaning of pedestrian exposure and its importance in pedestrian risk analysis.

The second chapter, "Area-Wide Methods," describes three general approaches to measuring pedestrian exposure for defined geographic areas, such as cities or counties. This chapter assists users in understanding the strengths and weakness of different methods of measuring pedestrian exposure over wide areas, and introduces users to existing sources of data on pedestrian activity.

The third chapter, "Site-Specific Methods," focuses on commonly used methods for counting pedestrian activity directly at specific sites, such as intersections or crossings. The performance of these methods is evaluated in terms of their relative cost, convenience, accuracy, and ability to collect a range of data points.

The fourth chapter, "Data Collection Planning at Intersections," assists users with the task of planning data collection at specific sites. It describes the statistical issues that must be addressed when designing a pedestrian data collection strategy, such as how to choose which sites to study and how to determine the number of sites to be studied.

The fifth chapter, "Estimating Annual Pedestrian Volumes," describes a method for converting short pedestrian counts into an annual measure of pedestrian volume using statistical analysis of pedestrian flow patterns. This method can be used to reduce the time and cost associated with developing an annual measure of pedestrian exposure, which is necessary to determine the annual pedestrian risk at a site.

Taken together, these chapters will assist the user in measuring pedestrian exposure for a variety of purposes and contexts. The purposes may include comparisons of the safety effects of pedestrian infrastructure; comparisons of pedestrian risk among different population groups; or comparisons of risk by mode of travel (e.g. walking versus bicycling). The geographic contexts may range from the entire state of California to a specific pedestrian crossing. Because each possible purpose and context will have a unique set of considerations and constraints, this protocol focuses on matching data collection methods with different study needs.

1.2. Who Should Use this Protocol

This protocol is intended to be used by traffic engineers and planners, consultants, and researchers interested in measuring pedestrian exposure. Although unaffiliated users will benefit from reading the protocol, it is most appropriate for those who are associated with an institution that has the resources necessary to mount a data collection program.

1.3. How to Use this Protocol

As discussed above, each chapter is aimed at a particular aspect of the data collection process. Some users may wish to read only the section that is most relevant to their needs. However, because the issues in the chapters are closely inter-related, many users will benefit from reading the entire document.

Users should understand that this protocol is not a "how-to" guide for measuring pedestrian exposure. Although many specific methods and equations are provided, the intention is to educate the user about the data collection process rather than to provide a set of instructions. This is because, as mentioned above, measuring pedestrian exposure is a complex task that is constrained by the study resources, purposes, and context. This protocol aims to inform the user about the data

collection strategies available to them, and to assist them in choosing which one best meets their needs.

2. PEDESTRIAN EXPOSURE

Before seeking to measure pedestrian exposure, it is important to have a clear understanding of the concept and its relationship to pedestrian risk. This chapter discusses the meaning of exposure in the context of risk analysis for pedestrian safety, and presents several common measures of pedestrian exposure used in the transportation safety field.

As this guide will demonstrate, there is no single best measure of pedestrian exposure, but some measures are better adapted to specific needs and purposes, such as comparing infrastructure; comparing risk among populations; or evaluating the change in pedestrian risk over time. This chapter will assist users in selecting an appropriate measure of exposure to match their needs.

2.1. Understanding Exposure and Risk

In epidemiology, exposure refers to a person's contact with a potentially hazardous situation or substance. For example, each time you fly in an airplane, you are exposed to ionizing radiation. Each time you cross a street, you are exposed to the possibility of being injured by a vehicle. Exposure can also be understood as a "trial event" in during which a harmful outcome might occur.

Risk is an abstract concept that refers to the probability a harmful event will occur given a certain number of trials. In pedestrian safety, each "trial" is a unit of exposure such as a minute spent walking or a road crossing Table 2.1 describes the relationship between exposure and risk.

Table 2.1: Exposure versus Risk

Exposure	Contact or amount of contact with potentially harmful situation (x)	(x)
Risk	Probability of collision/injury/fatality (c) per unit of exposure.	$P(c \mid x)$

The likelihood that any given trial event will result in a particular outcome is a function of the "chance set up". In transport safety, the "chance set up" is the transportation system itself, including its physical characteristics, users, and environment. Any one of these characteristics might influence the likelihood that a given trial event – such as a pedestrian crossing – will result in a collision (Hauer, 1982).

Risk and exposure are theoretical concepts that can only be indirectly estimated through the use of proxy measures. In the field of traffic safety, risk is typically represented by a simple ratio between collisions, injuries or fatalities, and exposure for a specific geography and time period (Chu, 2004). This ratio is referred to as the "collision rate" or the "accident rate". See Section 2.6 for a discussion of the limitations of collision rates as a proxy for risk.

Collision rate =	Number of collisions in a specified time and place	(1)
	Amount of exposure in a specified time and place	

If one finds that risk is higher at one intersection than another, it suggests that something in the "chance set up" (e.g. higher traffic speeds at one intersection) explains the difference. In this way, risk analysis is used to identify dangerous aspects of the transportation environment.

A short list of some of the factors thought to be associated with pedestrian risk include:

✓ Pedestrian characteristics including age and gender (Evans, 1991; Keall, 1995), and socioeconomic status and ethnicity (Ogden, 1997; Kraus et al., 1996). These characteristics may be related to distance and time traveled; pedestrian behavior; and awareness of the road environment.

✓ Pedestrian behavioral characteristics, such as risk-taking behavior, propensity to jaywalk, etc (Campbell et al., 2004).

✓ Trip characteristics: time of day/year, purpose, time elapsed between drinking alcohol and commencement of trip (Keall, 1995).

✓ Area characteristics related to transportation service and land use (Herms, 1970; Ossenbruggen, 1999).

✓ Roadway features such crosswalks and alternative crossing treatments, signalization, signing, pedestrian refuge islands, provisions for pedestrians with disabilities, bus stop location, and school crossing measures (Campbell et al., 2004).

2.2. Incorporating Exposure into Risk Measurement

Exposure is a crucial component of risk measurement. If the absolute number of injuries or fatalities is presented without controlling for exposure, it is easy to come to erroneous conclusions about risk.

The following graphs are provided to illustrate the importance of incorporating pedestrian exposure into measurement of risk. Figure 2.1 shows the number of pedestrians killed in New Zealand between 1988-1991, ordered by age and gender. These "raw" counts make it seem that children under twenty are most in danger of being killed.

However, when the raw counts are presented as a function of exposure, measured as the hours spent walking, a very different picture emerges (Figure 2.2). The age categories with the highest risk are those aged 80 and above and those ten and younger. Adolescents aged 15-20 do not have elevated risk levels; rather, the high numbers of fatalities in this category are due the fact that adolescents spend more time walking than other age groups.
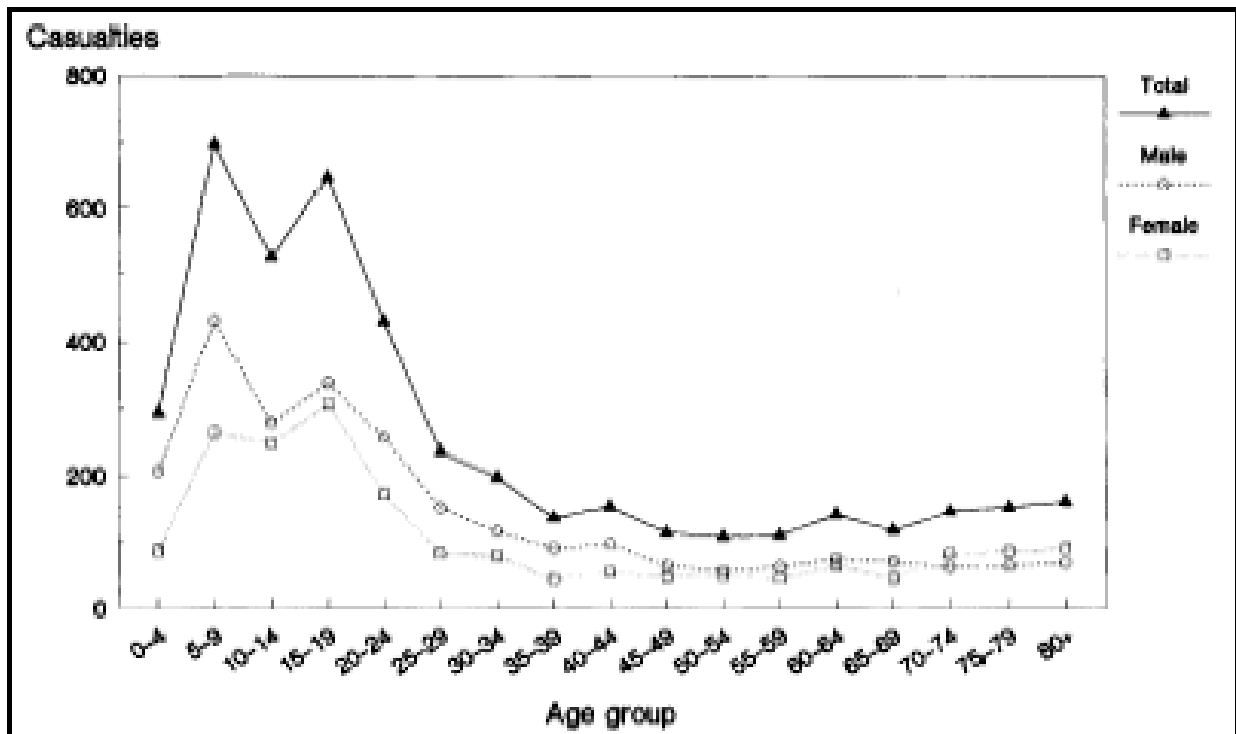
Figure 2.1: Number of pedestrians injured or killed in New Zealand, 1988-91 (Keall, 1995)
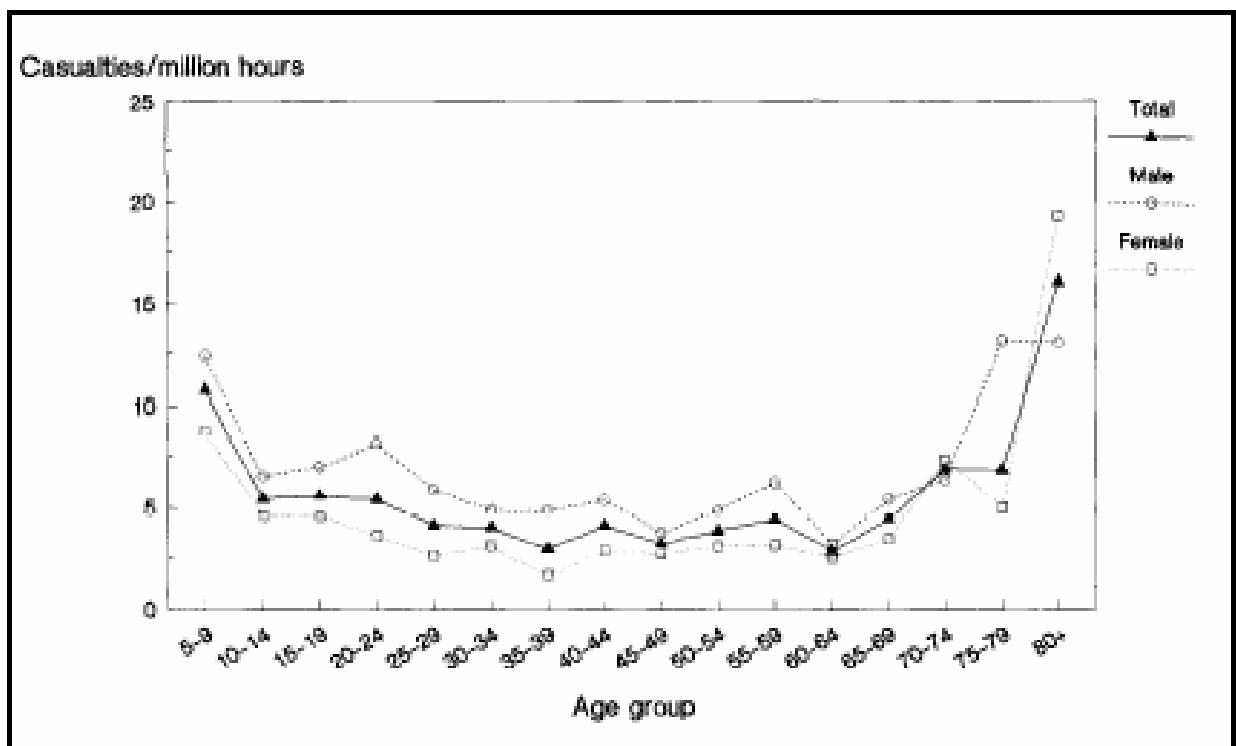
Figure 2.2: Number of pedestrian casualties per million hours walked in New Zealand 1988-91
(Keall, 1995)

When constructing a pedestrian safety risk measure, it is important to keep the following points in mind:

- ✓ The numerator and denominator in a risk measure must be consistent (Hauer, 2001); if exposure is in person-hours of pedestrian travel then the event in the numerator should be the number of pedestrians that experienced a collision or injury.

- ✓ The risk measure should reflect the type of risk being studied (Hakkert and Braimaister, 2002), such as whether the risk being studied is for an individual, or for a defined social group (Jorgensen, 1996).

- ✓ The denominator of the risk measure (pedestrian exposure) must reflect the intended purpose of the risk measure (Hakamies-Blomqvist, 1998). For example, a risk measure used to compare risk between different modes of travel should have a denominator (exposure measure) that is comparable across all modes.

- ✓ The denominator of the risk measure should reflect the target population being studied.

2.3. Defining Pedestrian Exposure

Pedestrian exposure is an abstract concept that reflects the opportunity for a potentially harmful pedestrian-vehicle interaction to occur; in other words, it is the number of trial events that could result in an injury or collision. It is very difficult to measure directly, since this would involve tracking the movements of all people at all times.

Instead, pedestrian exposure must be approximated using an appropriate proxy measure. Examples of measures used to represent pedestrian exposure at the micro level include pedestrian volume (Davis et al., 1988); the product of pedestrian and vehicle volumes at an intersection (Cameron, 1982) or roadway segment (Knoblauch et al., 1984); and the square root of that product (TRL, 2001). Measures used to represent exposure at the macro level in the U.S. include pedestrian distance traveled and pedestrian trips made (Pucher and Dijkstra, 2000, 2003); and the number of streets crossed (Roberts et al., 1996). In Europe, the most common

measures include the number of pedestrian trips made; time spent walking; and distance walked (ETSC, 1999).

In situations where travel-based measures of exposure are unavailable, population-based measures are sometimes used to approximate exposure (NHTSA, 2004). These may include population density (Qin and Ivan, 2001), and population divided by the percent of workers who reported that they usually walked to work in the last week (STPP, 2002, 2004).

The choice of exposure measure strongly impacts the resulting calculation of risk. For example, researchers at the Surface Transportation Policy Project used "miles traveled" as the denominator in estimating risk to pedestrians across the nation in the 2004 *Mean Streets* report. They concluded that walking is about twenty times more dangerous than riding in passenger cars, trucks, or on public transit (STPP, 2002, 2004). This conclusion can be distorted by the fact that walking is much slower per mile than other forms of transportation. If the researchers had used as the measure of exposure the amount of time spent traveling, rather than miles traveled, they may have reached different conclusions.

To illustrate further, Table 2.2 presents pedestrian collision rates in the European Union calculated using two different exposure measures: person-kilometers traveled and person-hours of travel. When person-kilometers walked is the measure of exposure, pedestrian travel appears to be many times riskier than travel by car. When person-hours spent walking is the exposure measure, then pedestrian travel appears to have the same risk as vehicle travel.

Table 2.2: Fatality Risks over Distance and Time for Travel Modes in the EU

Travel mode		108 person km	108 person hours
Road	Total	1.1	33
	Bus/Coach	0.08	2
	Car	0.8	30
	Foot	7.5	30
	Cycle	6.3	90
	M/C,MOPED	16.0	500
Trains		0.04	2
Ferries		0.33	10.5
Planes		0.08	36.5

Source: ETSC, 1999

2.4. Measures of Pedestrian Exposure

Presented in Table 2.3 is an exploration of some of the common ways that pedestrian exposure is measured. For each of these exposure measures, an explanation and examples are provided; common and appropriate uses are discussed; and benefits and limitations are explored. Not all possible ways of estimating pedestrian exposure are described.

Table 2.3: Common Metrics Used to Describe Pedestrian Exposure

	Explanation
Population	Number of residents of a given area, or number of people in a demographic group.
Number of pedestrians	Number of pedestrians observed in a given area during a fixed interval.
Trips	Number of distinct trips taken by an individual pedestrian.
Distance traveled	Total distance traveled by an individual pedestrian _or_ aggregate distance traveled by all pedestrians in a fixed area.
Time spent traveling	Total time traveled by an individual pedestrian _or_ aggregate time traveled by all pedestrians in a fixed area.

These examples will illustrate that there is no single best definition of pedestrian exposure. However, it is important to choose the definition of exposure that best matches the needs and purposes of the study. The chosen exposure measure should be compatible with the measurement devices being used and the target population being studied within a geographic area. The choice of exposure measure will also be determined in part by the amount of available resources, as some measures of exposure are more costly to collect than others.

2.4.1. Exposure based on population data

Population refers to the number of people who live in a given area, or the number of people who make up a particular demographic group. Because it is relatively easy and cheap to estimate, population data is often used as a simple proxy for pedestrian exposure.

There are a large number of issues that make the use of population highly unreliable as an exposure estimate. First of all, actual physical exposure to traffic is unlikely to be evenly distributed throughout the population. Second, time spent as pedestrians, or distance traveled, are not represented or accounted for in any way. Third, population does not necessarily relate directly to the actual number of people walking on the streets.

For example, some tourist sites attract a large number of people who are not accounted for by residential or employment population density, but who may still be involved in traffic collisions (Ivan et al., 2000). Models of pedestrian risk based on population provide only the roughest approximation, and are probably unreliable. Table 2.4 summarizes the issues related to exposure measures based on population.

Table 2.4: Exposure Based on Population Data

APPROPRIATE USES	✓ Used as an alternative to exposure data when cost constraints make collecting exposure data impractical ✓ Used to compare jurisdictions over time because population data is available for many geographies and time periods
HOW DATA IS GATHERED	✓ Population data for most cities is available on an annual basis through the American Community Survey (ACS). The ACS is administered by the U.S. Bureau of the Census and is accessible online (U. S. Census Bureau, 2006)
PROS	✓ Easy and low-cost to obtain; available for most geographies and time periods ✓ Adjusts for differences in the underlying resident population of an area – for example, sparsely populated suburbs versus densely populated inner-city areas ✓ Provides a crude adjustment for amount of vehicle traffic on the streets, since areas where more people live also tend to be areas where more people drive ✓ May be the only way to represent exposure if direct measurements cannot be taken
CONS	✓ Does not accurately represent pedestrian exposure ✓ Does not account for the number of people who travel as pedestrians in the area ✓ Does not provide information about amount of time or distance that members of the population were exposed to traffic
COMMON MEASURES	✓ Number of people in a given area: neighborhood, city, county, state or country ✓ Number of people in a particular demographic group: by age, sex, race, immigrant status or socioeconomic status
EXAMPLES	✓ In 2001, pedestrian collisions killed 20 people per million in California, but only 7 people per million in Nebraska. (FARS and U.S. Census data from 2001). ✓ In 2004, the male pedestrian fatality rate per 100,000 population in United States was 2.22, while the female pedestrian fatality rate was 0.95 per 100,000 population (NHTSA, 2004).

2.4.2. Exposure based on pedestrian volumes

Pedestrian exposure can be measured by the number of pedestrians that pass through a fixed point during a specified time interval. This is a common exposure metric, as it is relatively simple to assess through established manual and automated counting methods. This exposure measure is explained in more detail on Table 2.5.

Table 2.5: Exposure Based on Pedestrian Volume

APPROPRIATE USES	✓ Estimating pedestrian volume and risk in a specific location. ✓ Assessing changes in pedestrian volume or characteristics due to countermeasure implementation at that site.
HOW DATA IS GATHERED	✓ Manual or automated counts of pedestrians.
PROS	✓ Counts are simpler to collect than other measures such as time or distance walked. ✓ Automated methods for counting number of pedestrians are improving.
CONS	✓ Does not differentiate pedestrians by walking speed, age, or other factors that may influence individual risk. ✓ Does not account for the amount of time spent walking or the distance walked ✓ Not easily adapted to assess exposure over wide areas (for example, a city).
COMMON MEASURES	✓ Average number of pedestrians per day, sometimes called Average Annual Number of Pedestrians (Zeeger et al., 2005; Cameron , 1976; Hocherman et al., 1988) ✓ Number of pedestrians per time period, e.g., hour (Davis et al., 1988; Cove and Clark, 1993)
EXAMPLES	✓ The average daily pedestrian traffic at marked crossings was 312 pedestrians per site (Zeeger et al., 2005). ✓ Between 7:00 am and 10:00 am, 203 pedestrians crossed Rose Street at the intersection of Shattuck Avenue.

While the "number of pedestrians" is the term most frequently used to refer to this exposure variable, that terminology is not, strictly speaking, accurate. A more precise term is 'number of pedestrian crossings', since a single pedestrian can contribute to the count more than once if that person passes through the measurement point more than one time during the observation period (such as during an outbound journey, and then again on the return). In addition, it is important to distinguish whether the crossing is over a roadway or over an arbitrary line on a sidewalk. Statistics suggest that crossing the street might be more dangerous than walking along the road, so that crossing exposure should be distinguished from roadside or sidewalk exposure (Evans, 1991; Ossenbruggen, 1999).

Key to the accurate measurement of the number of pedestrians is a good operative definition of what constitutes an entry into the area, and what constitutes a pedestrian. For example, should a mother pushing an infant in a stroller be counted as one pedestrian, or two?

Any fixed point can be used. However, in practice, intersection crossings are often used as the fixed point. The reason for this is that crossing the street is an activity with a relatively high risk. In a study of pedestrian crash types across several states, Hunter et al. (1996) found that about a third of crashes involving a pedestrian occur at intersections, whereas only about 8 percent of all crashes occurred while the pedestrian was walking along the roadway.

A major assumption made in using an intersection as a fixed point is that each crossing represents a fixed unit of risk, independent of crossing distance or location within the crossing.

2.4.3. Exposure based on trips

Exposure based on number of trips estimates the number of walking trips taken by an individual, regardless of the distance or time the journey takes. Trips may be taken for the purpose of commuting to work or school, for social visiting, for utilitarian purposes such as shopping, for walking a dog, or walking purely for recreation. This information is generally gathered by surveying a representative subset of a population. Because other survey questions are usually asked at the same time, each trip can be linked to information regarding trip purpose, time of day, etc.

Number of trips as assessed by survey is usually difficult to relate to pedestrian collision data on a small-area scale. However, the data is useful to assess exposure over wide areas, especially when combined with other datasets, such as U.S. Census information or land use data, enabling additional analyses of factors affecting walking patterns.

Number of trips may not be the most useful metric for risk analysis purposes, but it is commonly used for assessing pedestrian behavior and activity, for making comparisons between large jurisdictions, and for examining changes over time (Table 2.6).

Table 2.6: Exposure Based on Trips

APPROPRIATE USES	✓ Assessing pedestrian behavior in large areas, such as cities, states, or countries. ✓ Examining changes in pedestrian behavior over time. ✓ Making comparisons between jurisdictions. ✓ Assessing common characteristics of walking trips, such as purpose, route, etc.
HOW DATA IS GATHERED	✓ Data is gathered through use of surveys, such as the National Household Travel Survey (2001)
PROS	✓ Appropriate for use in large areas. ✓ Best metric to assess relationship of walking with trip purpose ✓ Trips can be assessed as a function of person, household and location attributes.
CONS	✓ As with most surveys, a large number of respondents are needed to adequately represent the underlying population. ✓ Unlikely to provide information at the level of detail needed to assess risk at specific locations ✓ Pedestrian trips are often underreported in surveys (Schwartz and Porter, 2000)
COMMON MEASURES	✓ Average number of walking trips made by members of a population per day, week or year. ✓ Proportion of walking trips taken for particular purposes, such as commuting or shopping.
EXAMPLES	✓ In US, the percentage of all work trips made by walking fell from 10.3% in 1960 to only 2.9% in 2000 (Pucher and Dijkstra, 2003). ✓ While in the Mid-Atlantic States 15.8% of all trips are made by the walking mode, in the East South Central and West South Central states this percentage is around 6% (Pucher and Renne, 2003). ✓ In US, 38% of all pedestrian trips are made for social and recreational purposes and 32% for going to school and church, while 10% represent work trips (Pucher and Renne, 2003).

2.4.4. Exposure based on distance

Exposure based on distance, or distance traveled, represents the distance that pedestrians walk while exposed to vehicular traffic. This exposure measurement can be assessed on the level of the individual or on the level of the geographic area. On the individual level, exposure based on distance is expressed as the total or average distance that an individual pedestrian travels in a fixed time period, such as a day, week, or year. Typically the risk is stated in terms of the number of deaths per 100 million person miles traveled (Chu, 2003). As with the measurement of number of trips, assessment of this exposure measure is carried out through surveys of a

representative sample of the population. It is also possible to attach walking measurement devices, such as pedometers, to a sample of pedestrians.

On the geographic level, distance traveled is measured directly by aggregating the pedestrian distance traveled within a defined area during a fixed time period. This version of distance traveled is defined as the number of pedestrians counted, multiplied by the distance across the intersection. In this instance, the focus is on the total pedestrian-miles traveled, not the number of unique individuals traveling, and each individual may contribute distance more than once, if they pass through the observation area more than one time.

Using exposure based on distance to estimate risk, through either of the methods presented above, relies on the assumption that risk is a function of distance traveled. That means that other things being equal, crossing a roadway with four lanes carries twice the risk of crossing a roadway with two lanes.

The metric does not differentiate in terms of walking speed or other factors that could moderate the risk associated with distance. This potentially distorts the risk associated with walking when compared to other modes. One person-mile of walking represents far more exposure to vehicle traffic than one person-mile of riding in a passenger vehicle because of the differences in travel speeds between the modes (Chu, 2003). Thus, using a distance-based measure of exposure when comparing risk between modes may distort the results of the comparison. Table 2.7 presents more details about exposure measure based on distance.

2.4.5. Exposure based on time

Time exposure data has long been used for measuring risk (Jonah and Engel 1983; Anderson et al., 1989; ETSC, 1999). It has also been used to compare risk in different social groups or between travel modes. Keall (1995) estimated the risks of traffic collision for different sex and age groups by combining road collision data with survey data using the exposure measures "time spent walking" and "number of roads crossed". Chu (2003) proposed a time-based comparative approach to examining the fatality risk of walking and vehicle travel because time-based measures take into account the speed differences between walking and riding in a passenger vehicle.

Exposure based on time incorporates not only the distance traveled, but also adjusts for walking speed. Like distance traveled, time traveled can be measured on the individual level through surveys or through direct measurement at specific locations.

Time spent walking at a crossing, for example, might be measured by multiplying the number of pedestrians by the average crossing time. It can also be measured by adding the crossing times of each individual. In comparing two individuals, all other characteristics being equal, the measure will account for different walking speeds.

To better characterize the exposure measure based on time, Table 2.8 presents its appropriate uses and examples.

Table 2.7: Exposure Based on Distance

APPROPRIATE USES	✓ Estimating exposure at the micro or macro level. ✓ Estimating whether risk increases in a linear manner with distance traveled. ✓ Assessing how crossing distance affects risk
HOW DATA IS GATHERED	✓ For individual level exposure, through surveys such as the National Household Travel Survey (2001) ✓ For aggregate level exposure, measurement of the length of the area of interest, combined with a manual or automatic count of the number of pedestrians.
PROS	✓ Can be used to measure exposure at the micro and macro levels ✓ More detailed than pedestrian volumes or population data ✓ Can be used to compare risk between different travel modes ✓ Common measure of vehicle exposure
CONS	✓ Does not take into account the speed of travel and thus cannot be reliably used to compare risk between different modes (e.g. walking and driving) ✓ Assumes risk is equal over the distance walked ✓ Must typically assume that each pedestrian walks the same distance in a crossing or along a sidewalk
COMMON MEASURES	✓ Average miles walked, per person, per day. ✓ Total aggregate distance of pedestrian travel across an intersection.
EXAMPLES	✓ The 2001 fatality rate per 100 million miles traveled in the U.S. was 1.3 for drivers and their passengers and 20.1 for pedestrians (STPP, 2004). ✓ Between 1990 and 2000, the share of Americans walking to work fell from 3.9% to 2.9% (U. S. Census 2000 Summary File 3, Census 1990 Summary Tape File 3.)

Table 2.8: Exposure Based on Time

APPROPRIATE USES	✓ Estimating total pedestrian time exposure for specific locations. ✓ Comparing risks between different modes of travel (e.g. walking vs. riding in a car). ✓ Estimating whether risk increases in a linear manner with walking time. ✓ Comparing risk between intersections with different crossing distances and between individuals with different walking speeds.
HOW DATA IS GATHERED	✓ The number of persons passing through an area multiplied by the time traveled. ✓ Time spent on walking activities reported on surveys.
PROS	✓ Accounts for different walking speeds ✓ Allows for accurate comparison between different modes of travel. ✓ Can be used to measure exposure at the micro and macro levels ✓ More detailed than pedestrian volumes or population data
CONS	✓ Time based measures assume risk is equal over the entire distance of a crossing. Only a small portion of time spent walking on roadways represents real exposure to vehicle traffic. This portion would include time spent crossing roads, walking on the road surface, or possibly walking along the roadside where there are no curved sidewalks (Chu, 2003). ✓ Time spent on walking can be over estimated in surveys, because people perceive that they spend more time walking than they actually do (Chu, 2003). ✓ Walking may also be under-reported in surveys, because people may forget walk trips or may purposely choosing not to report. Both of these reasons are related to the fact that walking trips are relatively short. These very short trips may not register in the memory of respondents or the respondents may think that these short trips are unimportant (Chu, 2003)
COMMON MEASURES	✓ Average time walked, per person, per day or year. ✓ Total aggregate travel time of pedestrian travel across an intersection.
EXAMPLES	✓ In 2001, the U.S. annual per capita minutes traveled was 2,139 minutes (Chu, 2003).

2.5. Choosing an Appropriate Exposure Measure

Exposure can be estimated in a number of different ways for almost any situation, as summarized in Table 2.3. These different ways of assessing exposure lead to different risk estimates, each of which may be correct but each may convey a different meaning. When determining the best exposure measure for a given purpose, key considerations include:

✓ **What is the chosen method of measuring exposure? Does it match the study purpose?** Surveys will yield individual-level measures of exposure such as person-trips or person-distance walked, while direct observation will yield

geographic-level measures of exposure such as number of crossings or distance walked within a defined area.

✓ **Where is the exposure to be measured?** If exposure is measured at a facility such as a pedestrian crossing or along a sidewalk, then the exposure measure should be a micro-level measure, such as number of crossings.

✓ **What are the study resources?** Some exposure measures, such as time and distance, more accurately portray pedestrian risk than pedestrian counts alone. However, time or distance spent as a pedestrian will likely be more costly to collect than simpler measures of exposure.

The following section lists examples of study purposes and provides guidance on the choice of exposure measure for each.

2.5.1. Comparing safety infrastructure and countermeasures

When comparing the effects of infrastructure and/or countermeasure on pedestrian risk, the ideal measure of exposure will be collected directly in the area where the infrastructure and/or countermeasure are in place. This will allow an objective connection to be established between the site and pedestrian risk, and will allow a consistent numerator and denominator in the pedestrian risk measure. That is, the numerator will reflect the number of pedestrian-vehicle incidents occurring at the specific site and the denominator will reflect the number of "trials" occurring in the vicinity of the countermeasure. It should be noted however that surveys can in theory be used to track pedestrian use of infrastructure, although they are not well-adapted for this purpose. For example, the New Zealand Travel Survey of 1988-89 asked respondents to keep a diary recording the number of crossings made at 'zebra-style' pedestrian crossings (Keall, 1995).

The exposure measure should also be appropriate to the type of infrastructure being studied. If the effect of enhanced crossing devices is being studied, than the pedestrian crossing is an appropriate measure of exposure. Zeeger et al. (2005), for example, used the number of pedestrian crossings as the unit of exposure in a study comparing risk at marked and unmarked crossings. If the effect of new sidewalks

along the length of a block are being studied, then pedestrian distance walked along the block would be a better measure of exposure.

2.5.2. Compare risk between groups of pedestrians

If the purpose of the study is to compare risk among different groups of pedestrians, the measure of exposure should be linked to individual-level attributes such as age; racial or ethnic group; income category; and so on. For example, Keall (1995) estimated the risks of collision for different sex and age groups by combining road collision data with survey data using the exposure measures "time spent walking" and "number of roads crossed". These attributes are most easily collected through surveys, although it is possible to estimate certain pedestrian characteristics such as age and gender through direct observation.

2.5.3. Compare risk among different modes of travel

When comparing risk among different modes of travel, the best exposure measure reflects the different travel speeds of the modes being compared. For that reason, it is best to use time spent traveling to compare risk among different travel modes.

Because different modes use different infrastructure, it may be difficult to record and compare geographic-level measures of time spent traveling by various modes such as automobiles, airplanes, bicycles, and pedestrians. Recording the individual-level use of these modes by survey is more commonly used to compare risk.

2.6. Collision Rates as a Proxy for Risk

Although an in-depth discussion of risk measurement is outside the scope of this paper, it is important to be aware of possible pitfalls associated with using exposure data in simplistic risk analysis.

As noted above, exposure data is commonly used to calculate collision rates, namely the number of collisions in a given time and place divided by an exposure measure. The calculation of collision rates rests on the assumption that the number of collisions is proportional to exposure. In other words, it assumes that, all other things being equal, a place with more pedestrians should have more pedestrian-vehicle

collisions, and that the number of collisions should increase at a constant rate as the number of pedestrians increases. Figure 2.3 illustrates this assumption.
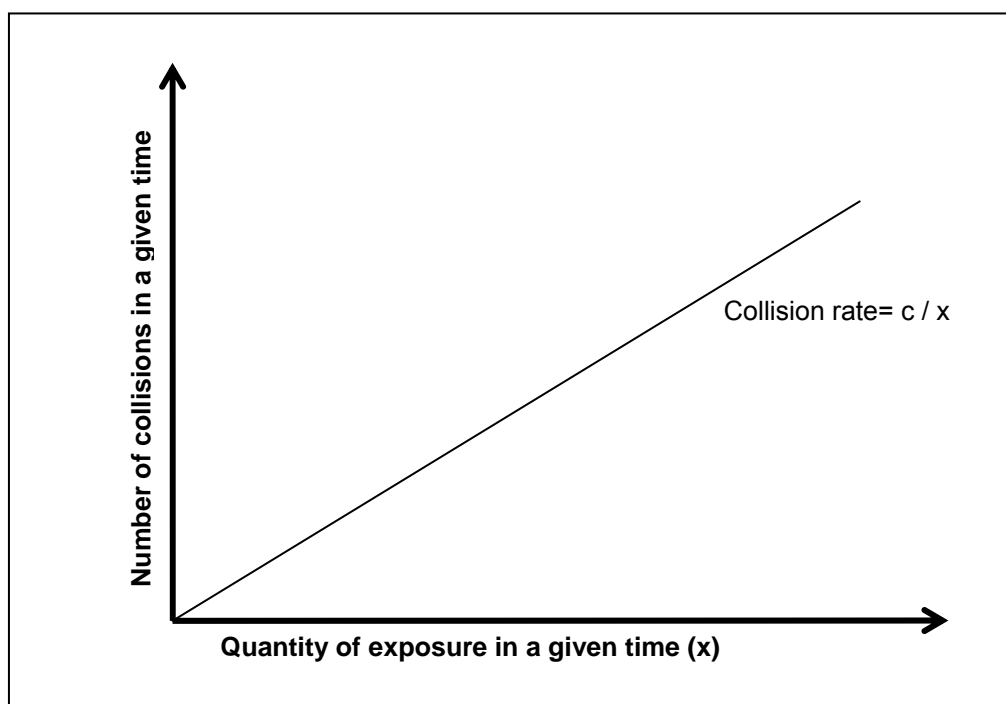
Figure 2.3: Assumed relationship between exposure and number of collisions

Although the assumption that collisions increase as a linear function of exposure is commonly made, there is substantial evidence to suggest that it is erroneous. Jacobsen (2003) has shown that pedestrian-vehicle collisions vary non-linearly with the number of pedestrians. In other words, risk appears to drop off when more pedestrians are present. Similarly, Lee and Abdel-Aty (2005) showed that pedestrian-vehicle collisions vary non-linearly with vehicle volumes. Collisions increase when more vehicles are present, but the rate of increase declines at high traffic volumes. The non-linear relationship may be due to more cautious driver behavior or reduced speed when many road users are present.

The calculation of collision rates without taking into account the non-linear relationship between exposure and collisions can lead to spurious conclusions in safety studies.

Hauer (1995) illustrated the pitfalls of collision rates using the following diagram (Figure 2.4). Accidents increase with exposure, but the rate of increase is not constant. The resulting curve is referred to as the "Safety Performance Function" of

the roadway. It may be empirically measured over time with the collection of accident data in periods of differing exposure.

Hauer (1995) shows how the collision rate (the slope of the curve) at point "B" in the diagram is lower than that at point "A" simply by virtue of the fact that the exposure has risen from 3,000 to 4,000 vehicles. If this fact is not taken into account, one could incorrectly conclude that a safety countermeasure was the cause of the decline in accident rates, when a change in exposure was alone responsible.
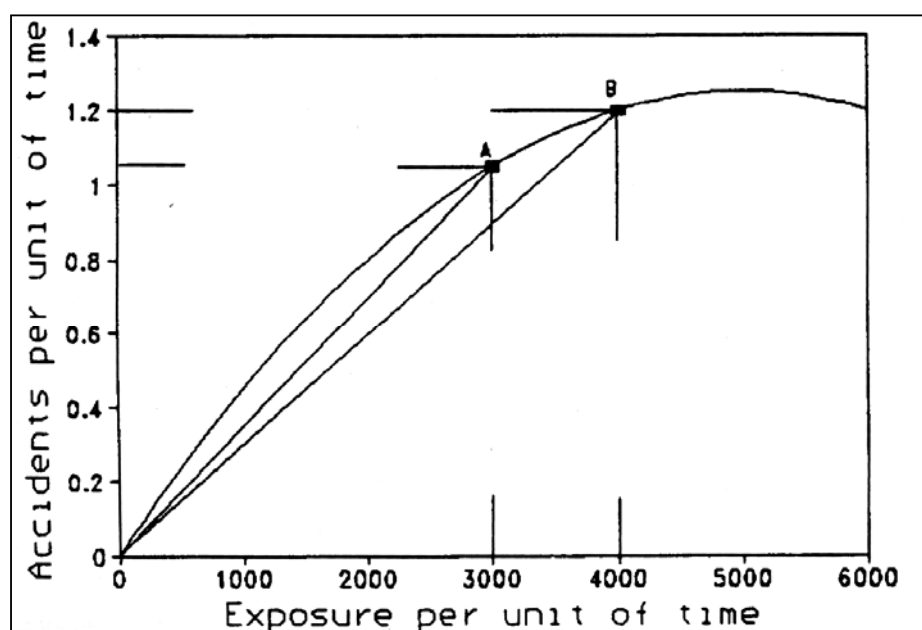
Figure 2.4 Non-Linear Relationship Between Exposure and Accidents (Hauer, 1995)

The best method of coping with the problems of accident rates is to discard them in favor of more complex models of risk. However, since risk modeling is often too costly for practical applications, accident rates are likely to remain common currency. Given that fact, it is sufficient to be aware that the usefulness of accident rates in measuring risk may be undermined in situations where exposure has changed substantially. Future studies of the relationship between pedestrian volumes and collisions are needed to define typical safety performance functions for pedestrian collisions. This will help identify the level of pedestrian exposure associated with a decline in collision rates.

2.7. Converting Between Exposure Measures at Pedestrian Crossings

As noted above, study resources may constrain the choice of exposure measure. For example, in areas with large numbers of pedestrians, recording the actual time each pedestrian spends at a crossing will require multiple observers, whereas recording the pedestrian volume will require fewer observers. In many cases, however, the estimated time a pedestrian spends crossing a street will provide a better indication of exposure than will a simple volume measurement.

In these cases, it is possible to convert the pedestrian crossing volume into an estimate of the aggregate distance crossed or time spent crossing. This can be achieved through the following equations (1) and (2).

Ped distance traveled (feet) = no. of crossings * distance crossed (ft)   (2)

Ped time walked (seconds) = Ped distance traveled (ft) / 4 (ft/s)[1]   (3)

Transforming pedestrian volume into time spent traveling or distance traveled at a crossing should be conducted for estimation purposes only. It should not be considered the "true" time spent traveling for the following reasons.

- ✓ Pedestrian crossing speed is not static but varies by pedestrian age; gender; pedestrian compliance with intersection controls; weather conditions; and signal cycle length (Knoblauch et al., 1996). One study noted that as many as 19 percent of pedestrians actually run across the intersection (Fitzpatrick et al., 2006).

- ✓ Pedestrians crossing distance is not static because some pedestrians may cross at an angle or walk outside the painted crossing.

- ✓ Pedestrian crossing speed alone does not fully account for crossing time because pedestrians who wait for signals to change require a "startup" time of approximately 3 seconds to begin walking (Knoblauch et al., 1996).

It should also be noted that this conversion should only be attempted for constrained areas where pedestrian distance walked can be estimated with reasonable accuracy.

---

[1] Pedestrian speed as indicated in the Federal Highway Administration 2003 Manual on Uniform Traffic Control Devices with Revision 1 Incorporated, published 2004

Observing pedestrian distance walked along a roadway, for example, is prone to error because individual pedestrians can stop, change directions, or enter and exit buildings, thus changing their distance traveled.

3. AREA-WIDE METHODS

The previous chapter illustrated the fact that there are several possible definitions of pedestrian exposure, and that the definition used in any given study is, to some extent, a function of the measurement instrument and the geographic context. This report identifies two main geographic contexts where measurement of pedestrian exposure takes place: wide areas, such as neighborhoods, cities, or the state, and specific sites, such as intersections or pedestrian crossings. These contexts can overlap when pedestrian exposure at specific sites is sampled in order to estimate exposure over a wide area.

This chapter discusses three general approaches to estimating area-wide pedestrian volumes. The first strategy involves directly sampling pedestrian activity at a representative set of sites throughout an area. The second strategy involves using surveys to gauge how much individuals report having walked in a given area. Surveys of this kind have already been implemented in some metropolitan areas and on the state level in California. The third strategy involves using modeling techniques to estimate pedestrian volumes from a combination of direct counts, surveys, and secondary data. The strengths and weaknesses of each of the methods listed above are discussed, and examples of each are provided.

3.1. Direct Sampling

Direct samples of pedestrian volume can be used to estimate pedestrian activity over a wide area. To achieve this, it is necessary to develop a strategy to sample volumes systematically through time and space. A systematic sampling design could be used to develop an estimate of the average volume at intersections in an area, for example. An in-depth discussion of representative sampling methods may be found in chapter 5, "Data Collection Planning at Intersections."

The direct sampling approach to measuring area-wide pedestrian volumes has some distinct advantages. Direct measurements of pedestrian activity are based on real observations, rather than reported behaviors, so they avoid the problem of under-reporting of short pedestrian trips common to surveys (Schwartz and Porter, 2000). Direct measurements capture the activity of all pedestrians at the sampled site,

regardless of age or economic status, although they do not capture the rich demographic information typically included in surveys. Direct measurements allow the linkage of pedestrian activity to site-specific factors such as intersection design.

Despite these advantages, there are very few examples of direct measurement approaches. This may be because of the lack of good inventories of the pedestrian network, which are necessary to devise a sampling scheme. The Institute of Transportation Engineers Pedestrian and Bicycle Council, with the assistance of Alta Planning and Design, have attempted to implement a program of pedestrian volume sampling over wide areas. This effort, known as the National Pedestrian and Bicycle Documentation Project, aims to establish a nationally consistent methodology for performing pedestrian and bicycle counts; to promote the performance of counts on official counting days during the second week of September; and to input counts into a national database (Alta Planning and Design, 2006). The project has resulted in collection of pedestrian volumes in a few cities throughout the nation. However, since there is no spatial sampling scheme associated with the project, the resulting volumes cannot be used to estimate pedestrian volumes over wide areas. The likelihood that the project will generate systematic, routinely collected pedestrian counts is small given its voluntary nature.

The best example of direct volume sampling comes from outside the pedestrian realm. The Federal Highway Administration has developed a Traffic Monitoring Guide to aid states in the systematic sampling of vehicle volumes. The guide describes a method for sampling every roadway section at least once within a six-year period, and for converting a point-measure of volume (Average Daily Traffic) into a distance-based measure (Vehicle Miles Traveled) based on the length of the roadway segment (FHWA, 2001). Although many states use the methods in the Traffic Monitoring Guide, some states, such as California, use a combination of direct counts and modeling to estimate vehicle volumes (Caltrans, 2005).

3.2. Surveys

Unlike direct sampling methods, surveys conducted at the local, state, and national level are commonly used to quantify pedestrian activity over wide geographic area. Because surveys are able to capture detailed pedestrian characteristics and preferences, they are very useful for studying the pedestrian behavior of specific

groups. Surveys are also able to capture detailed trip characteristics such as the number and length of walking trips made by an individual.

In direct sampling, by contrast, it is very difficult to determine the origin and destination of each pedestrian trip, or to determine detailed pedestrian characteristics. However, surveys have certain weaknesses. Surveys do not generally link pedestrian activity to specific infrastructure, such as roadway or sidewalk width, so it is difficult to determine the relationship between infrastructure and pedestrian activity from surveys alone. It is also difficult to determine whether the walking trips reported in surveys were made in areas where the pedestrian was exposed to traffic. Lastly, walking trips are commonly underreported in surveys, because individuals do not always remember short walking trips (Schwartz and Porter, 2000). For example, individuals may not report walking to access transit as a separate trip.

Survey data is available for many different types of geographies and time periods. When seeking information about pedestrian exposure over a wide area, it is important to know whether relevant survey data has already been collected. For that reason, this section focuses on describing existing pedestrian-related surveys and the type of information available from each. Three types of existing surveys are identified and evaluated: (i) health-related surveys; (ii) travel surveys; and (iii) the Journey-to-Work portion of the U.S. Census. These characteristics are also summarized in Table 3.2.

There will be cases where existing surveys will not always meet the data needs of the user. For example, there is no existing data source that provides an estimate of pedestrian exposure for the state of California as a whole on a frequent basis. In these cases, institutional support and resources are needed to implement more frequent or new data collection efforts.

3.2.1. Health-Related Surveys

Health surveys aim to track health conditions and risky behaviors. Since walking is a form of physical activity, some of these surveys include walking-related questions, which tend to be focused on whether the respondent obtained a healthy amount of physical activity. Therefore, these types of surveys may not contain information on

they exact amount of walking or whether walking took place in areas where pedestrians were exposed to traffic.

For example, the California Department of Health Services and the California Department of Transportation sponsored the Pedestrian Characteristics in California Survey in 2003 in order to track health trends. The survey included a question on the amount of time spent walking in a typical week (Schneider et al., 2005). Because the survey is not conducted on a regular basis, it is limited in its ability to track pedestrian volume trends over time, and it does not provide information about the total amount of exposure to traffic.

The Behavioral Risk Factor Surveillance System (BRFSS), an annual telephone survey administered by the Centers for Disease Control, is conducted annually. It includes questions on physical activity, but does not distinguish between walking and other forms of physical activity (BRFSS, 2006). The state of California could choose to add additional questions to the BRFSS in order to gain information about the prevalence of walking in the state.

3.2.2. Travel Surveys

Travel surveys are conducted at the metropolitan, state, and national level for transportation planning purposes. Most rely on travel diaries, in which respondents record detailed information about trips taken during a designated travel period. The detail provided by travel diaries is valuable in estimating pedestrian volume, because it allows volume to be expressed in terms of the amount of time walked, the distance walked, or the number of walking trips made.

The largest travel survey conducted nationally is the National Household Travel Survey (NHTS). The survey is conducted about every six years by the Federal Highway Administration, and records the travel patterns of about 20,000 randomly selected U.S. households. The NHTS reports the number of trips by mode that respondents took in the week the survey was administered. It can be used to quantify pedestrian trips as a share of all trips taken nationally or by major Census division (e.g. Mountain; Pacific, West South Central, etc.). The NHTS is not intended for use at the state or sub-state levels, but states or metropolitan areas can purchase add-ons (NHTS, 2006).

Several states and metropolitan areas also conduct travel surveys to serve local needs (TRB, 2006). In the state of California, travel surveys are conducted in several metropolitan areas and on at the state level. The California Statewide Household Travel Survey (CSTS), a travel survey of 17,040 California households, was conducted between 2000-2001 by the California Department of Transportation (Caltrans). The CSTS quantifies the number, duration, and approximate distance of trips taken by survey respondents on an average weekday for each mode of transportation. It also captures household demographic and economic characteristics.

The CSTS provides a robust estimate of the amount of pedestrian activity in the state of California, and for 17 sub-state regions, for the year 2000. The survey must be used cautiously or not at all for small geographic areas such as cities or counties (Caltrans, 2002). In addition, the CSTS cannot be used to track short-term trends in pedestrian activity because it is not conducted on a regular basis.

Several metropolitan areas in California also collect travel surveys similar to the CSTS and the NHTS. For example, the Metropolitan Transportation Commission conducts the Bay Area Travel Survey (BATS) a study of the travel patterns of approximately 15,000 Households in the 9-county Bay Area. The BATS was conducted in 2000, 1996, 1990, 1981, and 1965. The Sacramento Area Council of Governments and the Southern California Association of Governments also conduct travel surveys about once a decade.

3.2.3. U.S. Census Journey-to-Work and the American Community Survey

The Journey-to-Work component of the U.S. Decennial Census long form contains detailed information about the work-trip characteristics of one in six U.S. households. Respondents are asked about the location of their workplace; their usual means of transportation to work; and the amount of time it usually took them to get to work. The data is free to the public, available online, and covers large and small geographies throughout the nation.

However, Journey-to-Work data has some limitations. The survey questionnaire asks only about which mode of transport the respondent used most frequently to commute to work in the previous week. By doing so, it accounts only for work trips, which

make up a minority of all walking trips (Komanoff and Roelofs, 1993), and for employed adults, who make up less than half of the population (U.S. Census Bureau, 2004). Moreover, the form asks how the respondent "usually" got to work, and thus does not capture occasional trips to work made by another mode. Neither does it account for walking trips made as a component of the work trip, such as trips to and from a bus stop. This is because the survey questionnaire asks the respondent to name only the mode they used for the majority of the distance of their trip (U.S. Census Bureau, 2005).

In spite of these weaknesses, Census Journey-to-Work data has been used as proxy for pedestrian exposure because it provides some information about how much people are walking in an area, and is often the only data on walking available at the level of the city. One widely-known report on pedestrian safety, which was published by the Surface Transportation Policy Project, used the percentage of people walking to work and population data from the Census to compare pedestrian risk in metropolitan areas across the nation (STPP, 2002, 2004).

The Census long form that provides Journey-to-Work data is currently being replaced by a new product called the American Community Survey (ACS). Although the information being collected in the ACS is the same as what was collected in the Census long form, the two surveys differ in important ways. The most important difference is that Journey-to-Work data will be available every year through the ACS, rather than once a decade. Another important difference lies in the sample design. Whereas the Census long form data was collected during a specific week in April, the ACS samples households on a rolling basis during each month of the year. This means that ACS data will reflect traveler behavior throughout the year rather than for a specific season. When fully implemented, the ACS will sample about 3 million, or 1 in 10, U.S. households annually.

ACS data are currently available for communities of 65,000 or more on a yearly basis. For smaller communities, it will take between several years to accumulate enough samples to provide data. Beginning in 2008, yearly estimates based on three year averages will be available for communities of 20,000 or more, and beginning in 2010, yearly estimates based on five-year averages will be available at the Census

tract and block group level A summary of ACS data availability is displayed in Table 3.1.

Table 3.1: Block Group Level A Summary of ACS Data Availability

Type of Data	Population Size of Area	Data for the Previous Year Released in the Summer of:							
		2003	2004	2005	2006	2007	2008	2009	2010+
Annual estimates	≥250,000								→
Annual estimates	≥65,000								→
3-year averages	≥20,000								→
5-year averages	Census Tract and Block Group*								→

Data reflect American Community Survey testing through 2004

* Census tracts are small, relatively permanent statistical subdivisions of a country averaging about 4,000 inhabitants. Census block groups generally contain between 600 and 3,000 people. The smallest geographic level for which data will be produced is the block group; the Census Bureau will not publish estimates for small numbers of people or areas if there is a probability that an individual can be identified.

Source: U.S. Census Bureau, 2006

Table 3.2: Characteristics of Existing Pedestrian Related Surveys

Survey	Walking Question	Geographies	Years available
Decennial Census	Usual mode to work	Census tract → nation	1980, 1990, 2000
American Community Survey	Usual mode to work	Census tract → nation	Every year after 2003*
Behavioral Risk Factor Surveillance System	None-possible add on	States, nation	Every year
National Household Travel Survey	Number, length, duration of walk trips	Census divisions, nation	Every 6 years: 1969, 1997, 1983, 1990, 1995, 2001
California State Travel Survey	Number, length, duration of walk trips	Caltrans Districts, state of California	Every 10 years
Metro Area Surveys	Number, length, duration of walk trips	SF, La & Sac metro area	Varies –about every 6-10 years

*ACS release schedule varies by geography; data at the census tract level not available until 2010

3.3. Modeling Methods

Mathematical models can be used to estimate pedestrian volumes by combining key assumptions with existing data. If properly calibrated and tested, models can be powerful tools in estimating pedestrian volumes when direct measurement is not feasible. The advantages and disadvantages of modeling depend to some degree on

the model itself, but in general, models have the potential to save time and resources without overly compromising accuracy.

Radford and Ragland (2006) identified three main types of models: sketch plan models, network analysis models, and microsimulation models. The strengths and weakness of each for measuring pedestrian exposure are presented below.

3.3.1. Sketch plan models

Sketch plan models use available data to estimate pedestrian volumes for regional or city-wide planning purposes. These models rely on known or estimated correlations between pedestrian activity and adjacent land uses, such as square feet of office or retail space, and/or indicators of transportation trip generation such as parking capacity, transit volumes, or traffic movements (Schwartz et al., 1999). Some of these models are not capable of producing pedestrian volumes, but rather produce a dimensionless indicator of pedestrian activity.

The city of Sacramento, California, recently used a sketch plan method developed by Fehr and Peers Transportation Consultants (2005) as part of its pedestrian master plan. The method inputs demographic, economic and land use variables associated with walking into Geographic Information Systems software to produce a dimensionless "pedestrian demand index" for each street segment in the city.

3.3.2. Network analysis models

Network analysis models are more complex than sketch plan models because they rely on a map or model of the pedestrian network. As a result, they are capable of estimating volumes for specific street segments and intersections over an entire city or neighborhood. Although the models vary in technique, most use a variation on the four-step modeling approach to generate and distribute trips based upon assumptions about the amount of walking trips in a study area and various route choice algorithms (Senevarante and Morall, 1986; Ben-Akiva and Lerman, 1985; McNally, 2000).

Radford and Ragland (2004) used a network analysis model, Space Syntax, to estimate pedestrian volumes on streets and intersections throughout Oakland, California. The model required input of a pedestrian route map derived from publicly

available Census TIGER/line GIS centerline road maps; population and employment data from the U.S. Census and the California Economic Census; and raw pedestrian count data needed to calibrate the model. The model produced reasonable estimates of city-wide pedestrian volume.

The Space Syntax model is also useful for estimating pedestrian flow along corridors. This is very helpful because direct measurement of flow along corridors is difficult. It may be achieved by dividing the road network into small segments, such as a block length, and assuming that flow along the segment is constant. This is not always a fair assumption because of the complexity of pedestrian movement. For example, if a pedestrian is counted at the end of a block, it is uncertain whether she has been traveling for the entire block or if she just exited a building. With vehicle volumes, by contrast, it is often assumed that any vehicle passing through a point has been traveling along the length of the segment (FHWA, 2001). Space Syntax provides an alternative method of estimating flow along many corridors with a small set of samples as input.

3.3.3. Microsimulation models

Microsimulation models use flow principles from physical science to model pedestrian behavior in confined spaces such as the interior of shopping malls or subway stations, on a single or small number of streets, or within building interiors. Microsimulation models provide highly accurate, detailed information about pedestrian movement, but require specialized software, knowledge and extensive data inputs (Radford and Ragland, 2006).

3.3.4. Comparison of modeling techniques

Table 3.3 presents a comparison of these approaches, highlighting their advantages and disadvantages for estimation of wide-area pedestrian volumes. This table was adapted from Radford and Ragland (2006). Each of the modeling approaches discussed in this paper is suited to a different scale of geographic analysis. Sketch plan models are best for broad regional or statewide analysis; network analysis models are appropriate for corridor, neighborhood, or urban area analysis; and microsimulation models are best for a single street or smaller area.

Relevant literature indicates that sketch plans have the most potential to be put into standard use for estimating pedestrian volume throughout the state. While less accurate than other types of models, sketch plans are relatively simple to use and make the most out of existing data sources. A simple, standardized sketch plan method would be an improvement over the current absence of volume estimation methods in many areas.

Microsimulation models are much too complex and costly to be practical beyond the level of the street or intersection. Network analysis models have been successfully used to estimate pedestrian volumes in most large urban areas, but may be impractical in many small cities and rural areas that lack staffing and resources to perform the GIS analysis and calibration necessary to complete the model.

Table 3.3: Comparison of Modeling Methods

	Scale of Application	Advantages	Disadvantages
Sketch Plan	Large scale (city, region, state)	Little data collection required; No specialized expertise needed; Quick estimations.	Aggregate level; Low accuracy.
Network Analysis	Urban and neighborhood level	Good detail; Reasonable accuracy; Limited data requirements; Useful for estimating pedestrian flows along corridors; Appropriate to urban volume analysis.	Model must be calibrated with pedestrian counts; Requires existing GIS data; Must be submitted to sensitivity test.
Microsimulation	Individual Streets or intersections	Highly accurate; Detailed; Allows visualization of pedestrian flow.	Complex; Steep learning curve; Significant initial data requirements.

3.4. Comparison of Methods

This chapter reviewed and evaluated three possible systematic approaches to measurement of pedestrian volumes over wide areas. The choice of area wide counting methods depends on budget constraints and data needs, and the

availability of existing data. No single approach is best, but each has strengths and weakness. These are summarized in Table 3.4.

Table 3.4: Comparison of Approaches to Pedestrian Volume Estimation

Approaches	Advantages	Disadvantages
Direct sampling methods	Based on real, not reported pedestrian activity; All pedestrians at each site are sampled; Pedestrian volumes linked to specific sites; If designed appropriately, data could be aggregated from small to large geographies.	Difficult to devise a sampling scheme; Need a good inventory of the pedestrian network; Would require significant manpower; No demographic or attitudinal information captured; No information on distance, length, or time walked.
Survey methods	Can capture demographic and household data; Can capture distance, length, and time walked; Existing surveys could be adapted / expanded.	Walk trips are consistently underreported in surveys; Difficult to link walking to specific infrastructure; Difficult to determine whether walking occurred in areas exposed to vehicle traffic.
Modeling methods	Make the most of available data; Dynamic and flexible; Potential for lowest cost.	Different models may be needed for different geographic areas; Output may be limited to dimensionless measure of pedestrian demand.

4. SITE SPECIFIC METHODS

The previous chapter discussed approaches to measuring pedestrian exposure over wide areas such as cities or states. In many cases it is necessary to collect pedestrian exposure data at specific sites such as intersections, pedestrian crossings, or along a city block. Site-specific measurement of pedestrian exposure is used to identify high collision locations; to evaluate how infrastructure influences pedestrian risk; or to track changes in risk over time at a specific site or sites.

There are three main methods of counting pedestrians at specific sites: (i) field observation (ii) video observation with manual review and (iii) automated methods. This chapter describes these methods and evaluates the strengths and weakness of each.

4.1. Pedestrian Counts at Specific Sites

Pedestrian volumes at specific sites are usually collected directly using either (i) manual counts taken by collectors in the field or through video observation, or (ii) automated counts using specialized equipment. Push button counters are also used to count pedestrians. However, because of their lack of accuracy relative to the other counting methods, push button counters were not reviewed in this protocol. It has been determined that only 35 percent of all pedestrians use push button devices when they are available (Zeeger et al., 1982).

Pedestrian counting methods differ in their cost, convenience, level of data detail, and accuracy. In order to select the most appropriate method for different conditions and study purposes, it is important to understand the strengths and weaknesses of each method.

4.1.1. Manual counting methods

Manual counting methods are frequently used to quantify all types of transportation activity, including vehicle, bicycle, and pedestrian volumes. Manual methods are the most frequently used method of counting pedestrians, particularly for studies that require small samples of data at specific locations, such as pedestrian crossings. The two most common manual counting methods used to measure pedestrian flows at crossings are:

✓ *Field observations:* in which pedestrians are observed in the field and counted by hand.

✓ *Video-recordings:* in which camera recordings of pedestrian crossings are taken and then processed through playback and manual recording.

Field observations are typically used for periods of less than a day. In this case, the normal intervals for counting are 5, 10, or 15 minutes. The counts are recorded with tally sheets, hand-held computers, or clickers. Tally sheets can include an individual line for each pedestrian and his or her characteristics and/or behavior can be recorded, although not all tally sheets are designed this way. Some include only boxes in which the number of pedestrians crossing within a certain time are recorded. An example of a field sheet used to count pedestrians and make inferences about their characteristics is provided in Appendix A. Hand-held computers (PDAs) are more frequently used to count and classify vehicle movements, but can also be used to collect information about pedestrian flow and movement directions.

Clickers, Figure 4.1, are appropriate in situations where there is no need to record individual pedestrian characteristics. They are also helpful in areas of high volume, where it is important that the observer have his or her eyes focused on the street. Schweizer (2005) found that a person can count about 2,000 to 4,000 pedestrians in an hour using clickers, and only half that amount without them. Using more than one clicker, the field observer can factor in the difference between males and females or the direction of movement. However, recording these characteristics would decrease the capacity of the field observer.

Manual-video recording uses cameras to record images of pedestrians which are later reviewed by an observer. The observer records the number of pedestrians as well as pedestrian characteristics and behavior, if needed. Detailed review of behaviors, or crowded pedestrian conditions, may require that the observer review the video in variable time (e.g. slowing and speeding the video as needed).   Specialized video-playback tools may be used to facilitate review of the videos. One such tool was developed by the Partners for Advanced Transit and Highway Research (PATH), and is depicted in Figure 4.3.

The central issues with the manual-video method of counting pedestrians are the need for a good camera angle and resolution (Figure 4.2) and the long time required to review the video tapes, estimated to be three times the tape length (Diogenes et al., 2007).

Figure 4.1: Field Observation using clickers (Schweizer, 2005)

Figure 4.2: Video-image camera angle and resolution
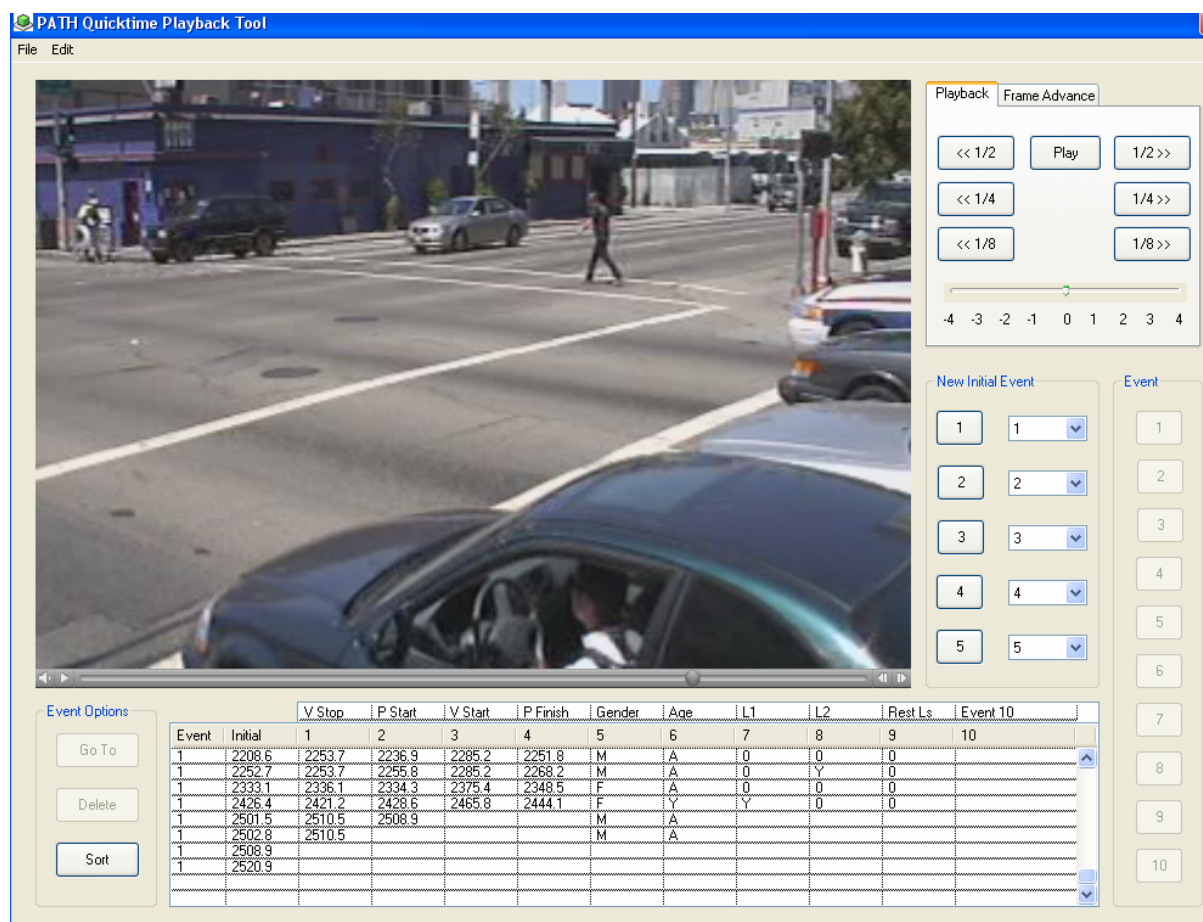
Figure 4.3: Path QuickTime Playback Tool

4.1.1.1. Cost of manual methods

The relative cost of field observations and video-recording counts varies based on the source of labor, the volume of the intersection, and the amount and type of data being collected. Costs can be broken into labor and equipment costs. In general, field observations are labor intensive but have low equipment costs. Video methods have higher equipment costs, and may have equally high if not higher labor costs depending on the amount of staff time taken reviewing video, and on whether the video camera can be left unattended in the field.

Cost of Manual Field Observations:

✓ Few equipment costs, though they may be increased if electronic hand-held devices (PDAs) are used to record pedestrian activity. However, use of these devices reduces the labor costs associated with data entry

✓ High labor costs. Staff are needed to observe pedestrians in the field and to perform data entry. More staff are needed at high-volume intersections, when several data points are being collected (e.g. pedestrian characteristics), or when detailed pedestrian behaviors are being investigated

✓ Training costs vary. The cost of training relates to whether consultant observers are used or whether observers are on staff, and to the need for data quality. Generally, more training can be expected to produce better quality data

✓ Costs can be reduced if counts performed by volunteers/students; if counts are integrated in to regularly scheduled vehicle counts (Schneider et al., 2005); and if counts are scheduled efficiently to maximize the use of available labor

✓ Example: the District Department of Transportation performs 10-hour counts at intersections across the city. The Department estimated in 2005 that each intersection counted cost between $400 - $500, including the cost of labor for pedestrian and motor vehicle counts and the cost of entering the field data into spreadsheets (Schneider et al., 2005)

Cost of Manual Video Recording:

✓ Equipment costs include the price of camcorders, tapes, and recording accessories. Camcorders vary in price depending on the quality required, but range from hundreds to a few thousand dollars. The cost of video tapes varies by number of hours recorded

✓ High-resolution or time-lapse video equipment may be required to record detailed pedestrian characteristics, or to monitor more than one crossing at a time. For example, the City of Davis, California, purchased a time lapse video system (including camera, playback system and videotapes) for $7,000 in 1998/99 (Schneider et al., 2005).The cost-burden of video equipment should be assessed over the life of the equipment

✓ Costs can be reduced if video counts are combined with other purposes, such as security

✓ Staff are needed for initial setup of camera and camera maintenance

✓ One staff person per video camera is typically required in the field to prevent vandalism and theft, unless the camera is concealed or made inaccessible

✓ Only one staff person is needed to review the video, regardless of the intersection volume, because video can be slowed down and rewound. However, staff may take many hours to review the video if detailed information or a high level of accuracy is required

✓ Transportation costs must be paid for staff and video camera. In some cases, a flat-bed truck may be required for set up of the video camera

4.1.1.2. Convenience and data detail of manual methods

Field observations and video-recording differ in their relative convenience and in the data detail that can be collected. Generally, field observers can capture a broad array of pedestrian characteristics and behaviors. Video-recordings are sometimes capable of capturing these details, but not without careful camera positioning and/or high resolution film. Video cameras may be able to record at times inconvenient for field observers, such as night time or weekends; however, this is only possible if the video is positioned or disguised such that it can be left alone without protection from vandalism or theft, and if the video image is unobscured by poor lighting or weather.

Convenience and Data Detail of Manual Field Observations:

✓ Staff schedules must be coordinated

✓ Inconvenient to collect data during inclement weather or during night/weekend hours

✓ Can waste labor time in areas of low volume

✓ Possible to capture detailed pedestrian characteristics like age, race, and specific behaviors (Mitman and Ragland, 2007; Diogenes et al., 2007)

✓ Difficult to record extra details if pedestrian volumes are high, unless additional staff are used

✓ Possible to capture mid-block crossings if observers trained properly

✓ Possible for a single staff person to observe multiple crossings if pedestrian volume is low

✓ Difficult to record the amount of time it takes pedestrians to cross

✓ Possible to record detailed information about the setting or nearby events that are not captured within a camera's field of view

Convenience and Data Detail of Manual Video Observations:

✓ If camera is positioned securely and disguised such that no on-site videographers are required to protect it, data can be collected at inconvenient times such as nights and weekends, assuming there is adequate lighting at the site

✓ If camera is rain-proof, it is possible to collect data during inclement weather

✓ Difficult to find a suitable place for video camera. Installation and use of cameras requires permits as well as security and safety procedures to protect the camera and those around it. For example, permits are typically needed to park a flat-bed truck near an intersection, and police must be notified so they do not suspect illegal activity.

✓ Difficult to capture pedestrian characteristics such as age or behavior without expensive cameras or precise positioning

✓ Presence of camera may influence pedestrian behaviors

✓ Cannot capture crossings from multiple directions unless multiple cameras are used or camera positioned at a very wide angle, which may compromise the image quality

✓ Cannot capture pedestrian behavior outside of the camera's field of view

✓ Possible to capture time and speed

✓ Cannot capture detailed information about the setting

4.1.1.3. Accuracy of manual methods

It is important to understand the accuracy of each counting method in order to make adjustments to counts or to choose the method with the desired level of accuracy. Although there are few empirical studies of the error of pedestrian counting methods, it is possible to identify and discuss the sources of error in each. In general, the accuracy of manual counts is affected by the level of observer training and attention, and whether the observer is in the field or reviewing video recordings. Mitman and Ragland (2007) compared the inter-reliability between different field observers and found there is a significant and measurable difference in the data quality produced by observers with different levels of motivation.

In both methods, error can be avoided by choosing observers carefully, conducting adequate training, and matching the collection method with location scenarios (Mitman and Ragland, 2007). However, video-recordings provide additional insurance against lack of observer motivation because they can be reviewed multiple times by different observers to check data quality.

Sources of error in manual field observations:

✓ Lack of attention. The motivation and training of field observers may affect their attention in the field.

✓ Differences in judgment. The unique personality attributes of field observers may affect their ability to judge pedestrian characteristics and behaviors, such as age and gender.

✓ Level of pedestrian activity. The amount of pedestrian activity may impact the accuracy of the count in a variety of ways. Very low or high volumes can impact the observer's attention and their ability to record all data points. More research is needed to determine the relationship between pedestrian volume and the accuracy of field observations.

✓ Amount of data needed. If it is necessary to record several data points for each pedestrian (e.g. gender, direction, age), the quality of the data recorded may decrease if the capacity of the observer is exceeded, or if recording the data requires the observer to take his or her eyes off the street.

✓ Length of time collecting data. If the collection period is long, the observer may take unscheduled breaks or get distracted.

Sources of error in manual video recordings:

✓ Lack of quality images. The camera angle, positioning, and image resolution affect the quality of the image and therefore the ability of the video observer to discern individual pedestrians and their characteristics.

✓ Differences in judgment. As with field observation, the attributes of video observers may affect their judgment of pedestrian characteristics.

✓ Lack of attention. As with field observation, the motivation and training of video observers may affect their attention. However, video recordings can be reviewed multiple times to ensure data quality.

✓ Traffic composition. Large vehicles may block the view of the crossings and render the video unusable in some instances. In contrast, field observers can adjust their viewing angle in real time to continue the observations and therefore eliminate this issue (Mitman and Ragland, 2007).

✓ Level of pedestrian activity. The level of pedestrian activity does not much affect the quality of counts because video can be reviewed in variable time to ensure all pedestrian are counted. However, the level of pedestrian activity may increase the time required to review the video, which may negatively impact the motivation of the video observer.

✓ Gaps in data collection. Data may be lost, and accuracy affected, when recording is stalled to change tapes, and if the camera malfunctions or is vandalized during counting.

4.1.2. Automated methods

In general, automated counting of pedestrians is advantageous because it can reduce the labor costs associated with manual methods. It also has the potential to record pedestrian activity for long periods of time that are currently difficult to capture through traditional methods.

Automated methods are commonly used to count motorized vehicles, but are not frequently used to count pedestrians at this time. This is because the automated technologies available to count pedestrians are not very developed, and their effectiveness has not been widely researched. Moreover, most automated methods are used primarily for the purpose of detecting, rather than counting, pedestrians (Dharmaraju et al., 2001; Noyce and Dharmaraju, 2002; Noyce et al., 2006).

A review of pedestrian detection technologies was performed by and Noyce and Dharmaraju (2002) and by Chan et al. (2006). Technologies include piezoelectric sensors, acoustic, active and passive infrared, ultrasonic sensors, microwave radar, laser scanners, video imaging (computer vision). A detailed review of these technologies and their potential for counting, not merely detecting pedestrians is being conducted for this project, and will be presented in the final report.

Of the technologies listed above, those most adaptable to the purpose of pedestrian counting are video imaging (computer vision) and passive infrared devices. Video imaging utilizes intelligent processing of digital images of pedestrians captured with a video camera (Figure 4.4) that is mounted above the area of pedestrian movement. The processor subtracts the static background from the image and then tracks the remaining objects to determine whether they are pedestrians (CLP, 2005).

Passive infra-red devices count pedestrians by tracking the heat emitted by moving objects. The company "Irysis", based in Great Britain, has developed infrared pedestrian counting devices that can be located either in or outdoors, and are mounted directly above the area of pedestrian activity (Figure 4.5). These sensors have the advantage of being relatively easy to install and configure, and are not affected by lighting conditions since they rely on heat to produce the images (CLP, 2005).

Figure 4.4: Video Imaging for Counting Pedestrian (CLP, 2005)
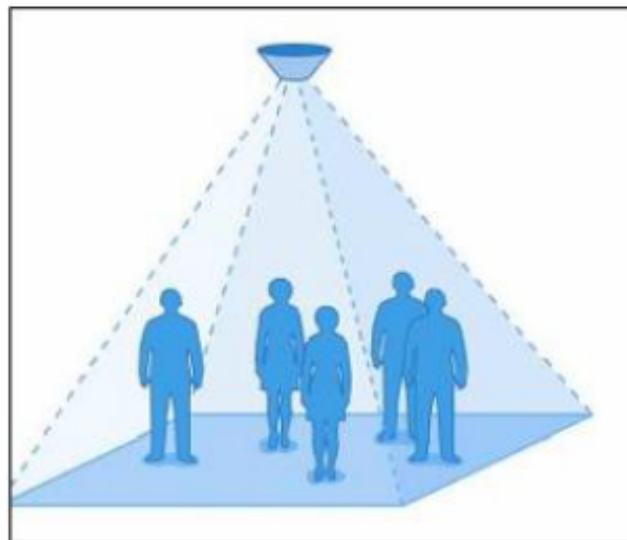
Figure 4.5: Irysis Infrared Pedestrian Counting Device (CLP, 2005)

4.2. Comparison Between Methods

The choice of pedestrian counting method should be based on the accuracy level desired, budget constraints, and the project data needs. For example, manual counts must be used when the effort and expense of automated equipment are not justified or when information about pedestrian characteristics or behavior is required.

To guide the selection of a method, it is important to review the advantages and disadvantages of each in collecting pedestrian exposure data at specific sites (Table 4.1). As specific advantages and disadvantages of the automated methods depend on the particular technology, only general aspects of these methods are highlighted.

It is important to emphasize that little is known about the relative accuracy and reliability of these methods. Field tests were performed within the context of this project to compare the particularities of the manual methods and a summary paper was submitted to the Transportation Research Board Conference (Appendix B). However, further work is needed to draw more specific conclusions about these methodologies.

Table 4.1: Comparison of Methods to Count Pedestrian at Crossings

Method	Advantages	Disadvantages
Field Observations	Relatively low cost; Observer can record detailed pedestrian characteristics and behaviors (Tally sheet)	Labor-intensive; Difficult to control the counting process; Problems at night, in unsafe locations, and during rainy weather; Cannot check accuracy of counts after they occurred;
Video Observations	Small error rate; Can replace several counters; Evaluation can be repeated several times; Possible to observe characteristics of road environment.	Difficult to find suitable place for video camera; May be gaps in the counting process (battery and tape change); Labor intensive (long analysis time) if good data quality is required; Can be hard to identify pedestrian characteristics and behaviors.
Automated Methods	Can collect data for long periods; Data storage is less time consuming.	Capital cost may be high; Specialized training may be required; Can not collect pedestrian characteristics / behavior.

5. DATA COLLECTION PLANNING AT INTERSECTIONS

Another aspect of site-specific measurement of pedestrian volume is the issue of where to collect data. The ideal would be to collect pedestrian volumes at all intersections of a city, but most projects have both budget and time constraints. In this case, a sample of the target population of sites must be selected for study. Nassirpour (2004) points out that there is no uniform standard of quality that must be reached by every sample and that the quality of the sample depends entirely on the stage of the research and how the information will be used. So, the development of a sample design that satisfies the project goals is crucial to obtain the necessary data efficiently.

This chapter describes a simplified set of statistical issues that should be considered when designing a methodology for collecting pedestrian volumes at intersections for different purposes. The proposed methodology is based on the recommendations of the Bureau of Transportation Statistics (BTS, 2003, 2005).

5.1. Sample Design Issues

Sample design is composed of three critical tasks: (i) definition of the target population; (ii) selection of sample technique; and (iii) determination of sample size. All these tasks have as constraints the objectives of the research, the type of the study and the resources available for the study, as shown in the Sampling Strategy Scheme of a Sampling Strategy (Figure 5.1). These constraints will play an important role when selecting the sample technique and determining the sample size.
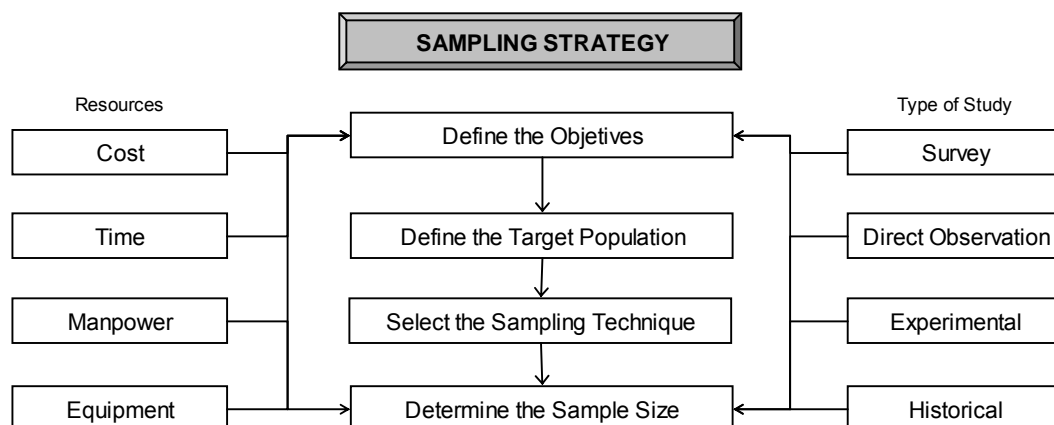
Figure 5.1: Generalized Model of Sampling (Adapted from Aggarwal, 1988 and Nassirpour, 2004)

5.1.1. Definition of target population

The target population can be defined as the complete set of sites from which you need to collect information (Nassirpour, 2004). Determining the population targeted is the first step in the sampling strategy and it is dependent on the study objective. For example, if you want to quantify pedestrian volume in the downtown's intersections, your target population is all the intersections in the downtown area. If you are interested in determining the average pedestrian volume in signalized intersections in California, so all signalized intersections within the state of California is your target population.

When defining the target population you must define the project objectives and specifications clearly to avoid collecting unnecessary data or generating bias. For example, if you want collect pedestrian volumes at marked and unmarked crosswalks you must define how to identify and distinguish between these intersections and define the geography of the study area.

After defining the target population, the operational sampling frame must be constructed. The sampling frame is a list of sampling units from which the sample can be selected at each sampling stage (Aggarwal, 1988). For example, in a study of intersection in the central business district, the sampling frame would be a database of all the intersections within the area. Ideally the target population must be coincident with the available list of sampling units. In situations where a complete database of the sampling units is unavailable, it is necessary to adjust the sample from the frame population to the target population.

In traffic observation studies, the Geographic Information Systems (GIS) and digital road databases are commonly used to develop the sampling frame (Shapiro et al., 2001). GIS can be very useful in defining the sets of intersections that are eligible for sampling, and can also provide additional information about the site, such as the number of pedestrian collisions.

5.1.2. Selection of sampling technique

After selecting the target population it is necessary to choose a sampling technique (Figure 5.2). The first step in selecting this technique is to decide whether to use non-probabilistic or probabilistic sampling.
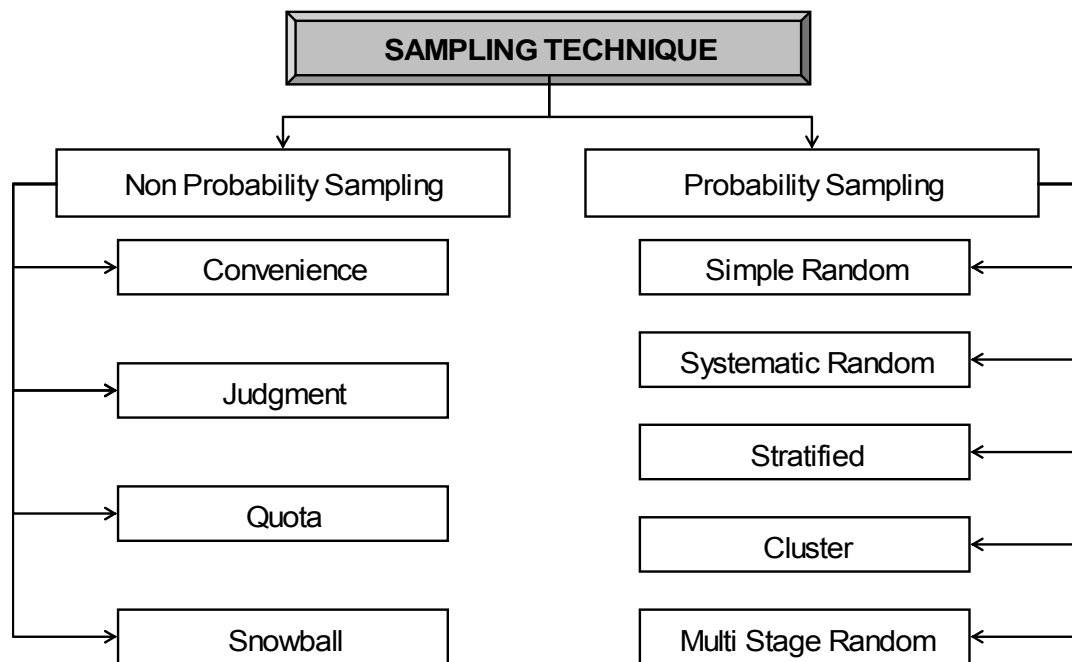
Figure 5.2: Classification of Sampling Techniques (Adapted from Aggarwal, 1988)

The non-probabilistic samples are selected through non-random methods, where the researcher has a lack of control over the sampling error. This type of sampling is most often used in experimental studies or case studies, when the researcher is interested in specific units or individuals and not in making conclusions about an entire population.

Non-probabilistic samples do not require the determination of sample size. Instead, the researcher will typically select a small number of samples based on subjective criteria. Table 5.1 describes in few words some of the existing non-probabilistic sampling techniques, pointing out the advantages and disadvantages of each method.

In contrast to non-probabilistic sampling, probabilistic sampling involves the use of statistical principles to select units or individuals randomly. This allows the researcher to calculate the sampling error and to make inferences about the target population. Probabilistic sampling requires more time and money to design the sample and to calculate the sample size necessary to obtain a representative sample. Table 5.2 describes the most frequently used probabilistic sampling techniques.

It is important to keep in mind that the selection of a sampling technique must be based on the research objectives and on the type of study.

Table 5.1: Non-Probabilistic Sampling Techniques

Non-probabilistic method	Definition	Example	Advantage	Disadvantage
Convenience	Obtaining a sample of people or units that are most convenient to study.	Selecting intersections with available collision data	Low Cost; Easy method of sample design.	No representative sample; Not recommended for descriptive or casual studies.
Judgment	Selecting a sample based on individual judgment about the desirable characteristics required of the sampling units.	Selecting signalized intersections because of experience or intuition that they have higher pedestrian flow.	Low cost; Allow to draw some conclusions about the characteristics of the selected sample.	Does not allow drawing general conclusions about the entire population.
Quota	It is similar to the judgment sample, but requires that the various subgroups in a population are represented.	Making sure to select some signalized and some unsignalized intersections in a sample.	Low cost; Allow to draw some conclusions about the characteristics of the selected sample.	Does not allow drawing general conclusions about the entire population, or sample subgroups.
Snowball	Additional survey respondents are obtained from information provided by the initial sample of respondents.	Used when surveying individuals about their behaviors (e.g. how much they walk in specific areas)	Some characteristics about the target population can be known	Requires a lot of time and resources; Used only for surveys.

Table 5.2: Probabilistic Sampling Techniques

Probabilistic method	Definition	Example	Advantage	Disadvantage
Simple Random	A sampling procedure that ensures each element in the population will have an equal chance of being included in the sample	When there are enough resources; to inquire about the characteristics of the entire population	Simple; Conclusions about the population can be drawn.	Subgroups within the target population may not be represented in the sample; Larger samples are necessary.
Systematic Random	Samples are randomly selected from a list in order, but not every one has an equal chance of being selected.			The sample may not be representative because of the ordering of the original list.
Stratified	Sub-samples are drawn within different strata. Each stratum is composed of samples with similar characteristics.	When representation of all subgroups within a particular sample is necessary.	More efficient sample (variance differs between the strata); Small sampling error between strata; Smaller samples.	May be difficult to determine characteristics of individuals to appropriate classify them in specific strata.
Cluster	Entire groups, not individuals, are selected to participate in the data collection; Simple random sampling is applied to the representative "clusters" to select the clusters in which all members will participate.	When the population is too big or when there is a lack of information about individual sampling units (e.g. all vehicle occupants in the United States)	Efficient for large numbers. Do not need to identify all units. Smaller samples; Less expensive relative to the population size.	Sample may not be as representative as desired; Error may be greater than with other techniques; Pilot studies may be necessary to identify the clusters.
Multi Stage Random	Stratification techniques within the clusters used to refine and improve the sample. Examples of this kind of sampling: National Safety Belt Survey.			Like cluster sampling but more representative within clusters.

* Based on Nassirpour, 2004 and MRUTC, 2005

5.1.3. Determination of sample size

There are many considerations that come into play when determining the sample size, such the level of precision to be achieved, operational constraints, available resources and the chosen sampling technique. The more accurate the desired results, the greater the sample size required. In order to achieve a certain level of precision, the sample size will depend, among other things, on the following factors (Statistics Canada, 2006):

✓ *The variability of the characteristics being observed*: If all intersections have the same pedestrian flow, then a volume count in one would be sufficient to estimate the average pedestrian flow for all the intersections. If intersections have very different flows, then a bigger sample is needed to produce a reliable estimate.

✓ *The sampling and estimation methods*: Not all sampling and estimation methods have the same level of efficiency. Operational constraints and the unavailability of an adequate frame sometimes mean that the most efficient technique cannot be used. A larger sample size is needed if the method used is inefficient.

Som (1996) points out other important observations about sample size:

✓ Estimates of sample size required to obtain measures with a given precision will often be found to be quite large, when derived on the basis of unrestricted simple random sampling;

✓ Small samples have proved useful, not only as pilot studies to full-scale surveys, but also providing interim estimates;

✓ An organizations with inadequate resources can start from a small sample and with increasing resources build up a fully adequate sample; the Current Population Survey of the U.S.A., for example, started in 1943 with 68 primary areas which were enlarged to the present 449.

✓ It is possible to combine smaller monthly or quarterly estimates into yearly estimates, and the yearly estimates into estimates covering longer periods, to provide estimates with acceptable precision.

✓ In the interest of true accuracy, it may sometimes be better to conduct a smaller sample with adequate control than try to canvass a much larger sample but with poor quality data.

In this protocol, examples are given on how to estimate the sample size for collecting pedestrian volumes at intersections for different purposes. However, these examples are based on specific scenarios, and if any variable of the scenario is changed the sample size must be recalculated.

5.2. Sampling Intersections in a City

As presented above, the sample design must be based on the research objective, the type of study and the available resources. Therefore, when planning to collect data about pedestrian exposure at intersections, the data needs and goals must be clearly defined. These considerations include: (i) what data items are needed and how they will be used; (ii) the precision level required for estimates; (iii) the format, level of detail, and types of tabulations and outputs; and (iv) when and how frequently users need the data (BTS, 2005).

Once data needs are defined, the existing data collection systems must be reviewed in order to determine whether all or part of the required data are already available, or could be more easily obtained by adding or modifying other data collection systems (BTS, 2005). Sometimes, manual pedestrian counts can be combined with existing motor vehicle counts at little or no additional cost. This has already been achieved with good results in some U.S. communities such as Albuquerque, NM, Baltimore, MD, and Washington, DC (Schneider et al., 2005). Pedestrian counts can also be combined with other initiatives such as general plans, pedestrian plans, or studies (e.g. the National Seat Belt Survey). When it is not possible to obtain the necessary pedestrian exposure data by adding or modifying the existing data collection system, a sample design is needed.

Data collection and analysis occurs after the data collection methodology has been defined. However, in systematic studies where data collection is performed repeatedly, it is necessary to reevaluate the study objectives and methodology each time data is collected, creating a loop in the data collection planning process. This

loop ensures changing conditions are reflected in the study design. Figure 5.1 illustrates this process.

```
                    ┌─────────────────┐
                    │   Define Goals  │◄──────────────────┐
                    └────────┬────────┘                   │
                             │                            │
                    ┌────────▼────────┐                   │
                    │ Determine Data  │                   │
                    │     Needs       │                   │
                    └────────┬────────┘                   │
                             │                            │
                 ┌───────────▼───────────┐                │
                 │  Could data be        │          Review the
                 │  obtained by adding   │          Initial Objectives
                 │  or modifying other   │                │
                 │  data collections     │                │
                 │  systems?             │                │
                 └───────────┬───────────┘                │
            YES    ┌─────────┴─────────┐   NO             │
          ┌────────▼──────┐       ┌────▼──────────┐       │
          │ Define        │       │  Develop a    │       │
          │ approches to  │       │  new data     │       │
          │ modify or     │       │  collection   │       │
          │ combine       │       │  system       │       │
          │ existing data │       └────┬──────────┘       │
          │ collection    │            │                  │
          │ systems       │            │                  │
          └────────┬──────┘            │                  │
                   │      ┌────────────▼────┐             │
                   └─────►│                 │             │
                          │   Collect data  │─────────────┘
                          │                 │
                          └─────────────────┘
```
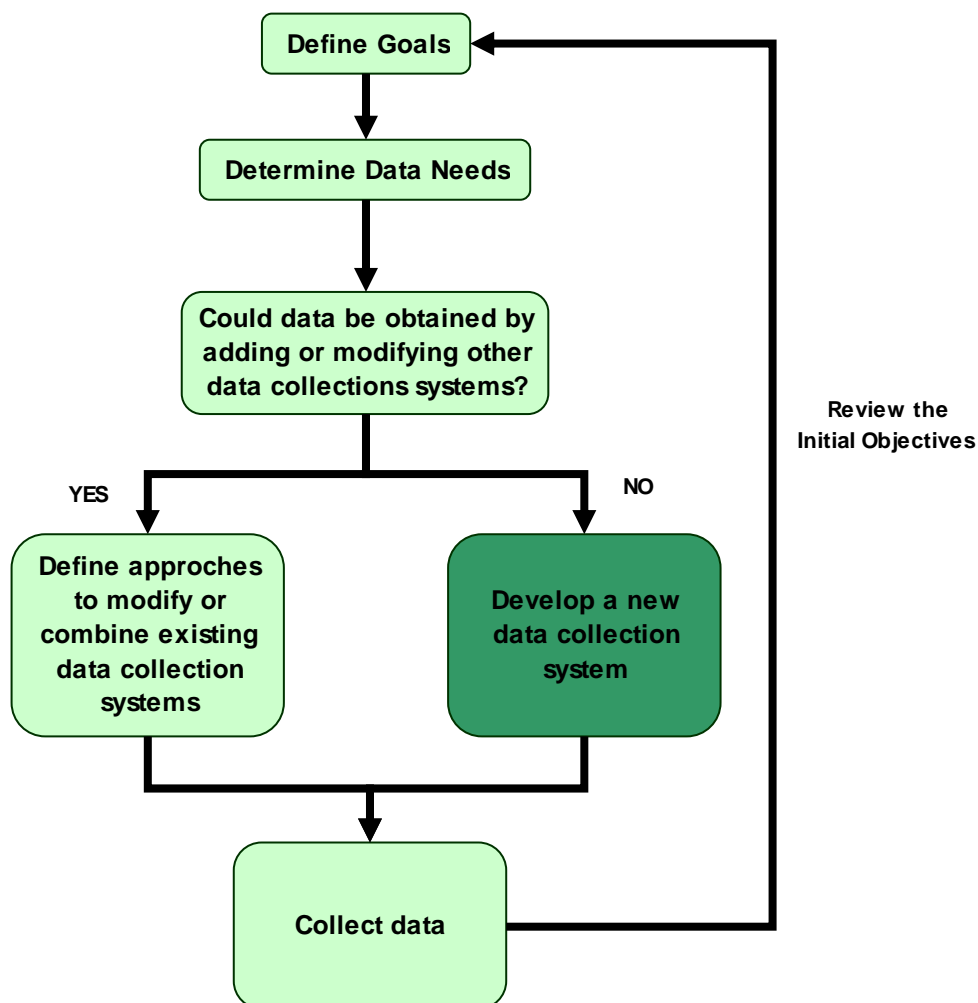
Figure 5.3: Methodology for Planning Pedestrian Exposure Data Collection at Intersections

This chapter focuses on the development of new data collection systems. Three hypothetical scenarios involving the collection of pedestrian exposure data were constructed to illustrate the necessary procedures. These scenarios are intended to be brief sketches of data collection planning. Not all methods and purposes are explored in the scenarios.

To simplify the analysis of the scenarios, we have organized the sampling design in 4 steps, as shown in the Figure 5.4.
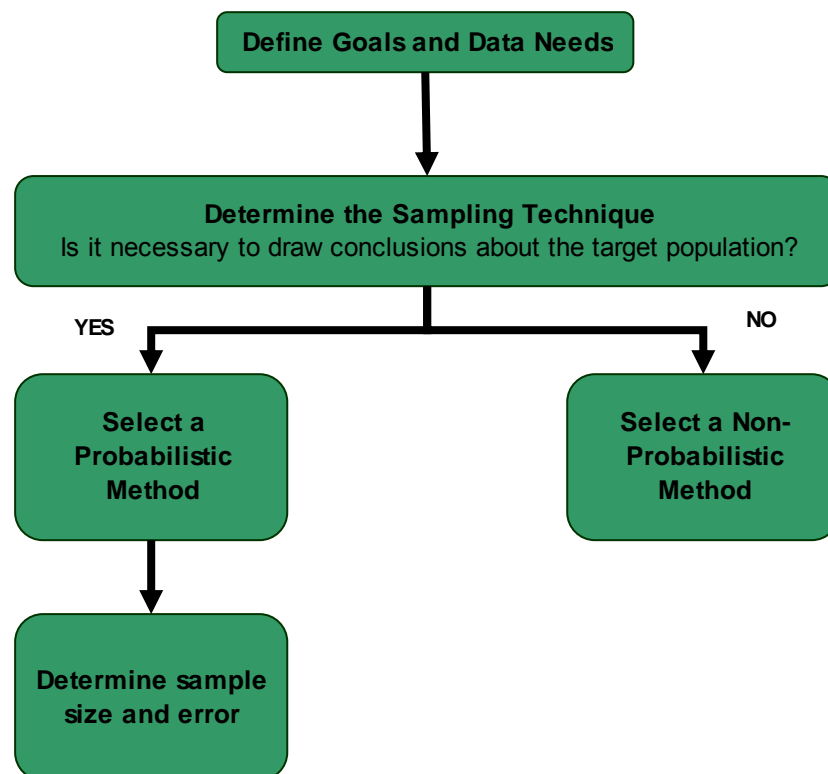
Figure 5.4: Sampling Design Steps for Pedestrian Exposure Data Collection at Intersections

5.2.1. Scenario 1: Evaluate change over time

One of the uses of pedestrian exposure data is to evaluate change over time, such as the change in pedestrian risk in an area or a countermeasure's effectiveness (before-and-after studies, such as Banerjee and Ragland, 2007). In such circumstances, it is common that the researcher is more interested in studying specific sites using non-probabilistic methods to choose where to collect data.

In the first scenario the research goal is to evaluate pedestrian risk among 10 specific intersections before and after signalization. In this case, there is no need to make general inferences about the sample population, and the sites are already chosen using the judgment method (i.e. the intersections that will be signalized). However, the researcher must be aware that when evaluating a temporal series of data it is important to use the same methodologies through time, thus avoiding seasonal influence (Cameron, 1976; Hocherman et al., 1988; Hottenstein et al., 1997).

5.2.2. Scenario 2: Evaluate risk related to infrastructure type

Pedestrian exposure can also be used to compare the safety associated with infrastructure. For example, Zeeger et al. (2005) compared pedestrian risk among marked and unmarked crosswalks. For this purpose, judgment samples or random samples can be used.

The research goal of the second scenario is to determine if pedestrian collision rates at marked mid-block crossings are higher than at unsignalized intersections. The available annual numbers of collisions are aggregated by type of crosswalk in business area of San Francisco. Therefore, the sample frame is marked mid-block crossings and unsignalized intersections in the San Francisco central business district.

To perform the analysis, the annual volume of pedestrians at each type of crossing must be determined. Since the study goal is to understand target population characteristics, a representative sample is needed.

Two random sample sites must be selected: one to determine the annual pedestrian volume at mid-block crossings and one to determine the annual pedestrian volume at unsignalized intersections.

Sites with similar characteristics are expected to have similar pedestrian flows, meaning that the variance in a sample is likely to be relatively low. In this case, a simple random sample technique is appropriate. It is very simple to apply when there is a complete list of all targeted crossings available, and will result in a small sample size when the variance between selected units is low.

Each sample size can be determined by the formula (3).

$$n = \frac{z^2 CV^2}{e^2} \qquad \textbf{(4)}$$

where,
z is the z value, which is derived from the desired confidence level (e.g., 1.645 for 90% confidence level, 1.96 for 95% confidence level, and 2.575 for 99% confidence level);

e is the margin of error (e.g., .07 = + or – 7%, .05 = + or – 5%, and .03 = + or – 3%); and

 CV is the coefficient of variance of an attribute in the population (e.g., .10 or 15% for moderate variances).

If a confidence level of 95% (z=1.96) is adopted, with the maximum acceptable error of 5% and a low coefficient of variance (10%) is assumed, the sample size must be 16 crosswalks for each type, totaling 32 intersections. After the first round of data collection the coefficient of variance must be calculated and the sample size must be estimated again, in order to optimize the sample size with a reliable and accurate sample.

The crosswalks must be sampled randomly in each subgroup (mid-block and intersection crossings). It is therefore necessary to have a complete list of all units of the target population classified by subgroup.

5.2.3. Scenario 3: Sampling exposure in a geographic area

Sometimes it is necessary to determine pedestrian exposure in certain area: (i) to compare pedestrian risk between different cities; or (ii) to estimate pedestrian risk for the area. In these cases, a probabilistic approach is necessary to be able to estimate the exposure measure accurately and a stratified sampling technique is most appropriate, since it can provide a sample representative of defined subgroups.

In the third scenario, the main objective is to asses pedestrian risk in the city of Berkeley systematically (the data collection must be repeated every 5 years). The estimate must be representative of the volumes at different types of intersections at different areas. So, a stratified sample must be designed.

Strata must be defined taking into account the similarity of intersection characteristics and geographic sub-areas. One can classify the intersection by type (signalized or non-signalized) or by function (Arterial/Arterial; Arterial/Collector; Arterial/Local; Arterial/Access Ramp; Collector/Collector, Local/Local). There are also many ways to classify geographic areas[2], but in this scenario they are defined in 3 categories: Central Business District; Fringe area; and Suburban and Rural

---

[2] Geographic area classification is explained in greater detail in Chapter 6.

Area. The number of strata will determine the sample size needed, as more strata will require a larger number of samples. For the first year of data collection, it is reasonable to simplify the data collection and use a small number of strata for each stratification variable.

In this scenario, the sample is divided in two stratification variables: (i) intersection type with two classes and geographic area with three classes. Table 5.3 presents these variables, which total six strata (3 x 2). To calculate the number of sites needed within each stratum, the same equation used for scenario 2 can be used (equation 3).

Table 5.3: Stratification Variables

Stratification Variable	N°. of classes	Classes Description
Intersection Type	2	Signalized
		Unsignalized
Geographic Area	3	Central Business District
		Fringe area
		Suburban and Rural

Adopting a confidence level of 95% (z=1.96), with the maximum acceptable error of 5% and assuming a low coefficient of variance (10%), 15.4 intersections must be selected within each stratum. Therefore, a minimum of 93 intersections (15.4x 6 strata) must be sampled. As in scenario 2, the true coefficient of variance must be calculated and the sample size must be reevaluated after the first round of data collection.

To obtain a more representative sample, we can distribute the total sample size among each stratum proportionally to the target population profile. For example, if in Berkeley 30% of intersections in the central business district are signalized, then 28 intersections with this characteristic must be randomly sampled. However, at least ten units within each stratum should be sampled to maintain statistical reliability.

6. ESTIMATING ANNUAL PEDESTRIAN VOLUMES

In order to determine the annual pedestrian collision risk at a specific site, two pieces of information are needed: the annual number of pedestrian collisions and the annual pedestrian exposure. The numerator of the risk measure, which is the annual number of pedestrian collisions at a site, can be obtained relatively easily from the Statewide Integrated Traffic Records System. The denominator of the risk measure, which is the annual pedestrian exposure at the site, is more difficult to obtain, because it is usually impractical to measure pedestrian volumes continuously for an entire year.

The process of estimating annual pedestrian exposure can be simplified using extrapolation techniques. These techniques allow short samples of pedestrian volume to be converted into a measure of annual pedestrian exposure. The purpose of this chapter is to describe a commonly used method of extrapolating pedestrian volumes and to provide examples of the application of the method.

6.1. Approaches to Estimating Pedestrian Volumes

In theory, the annual pedestrian volume at a site can be obtained by observing and recording pedestrian flow continuously throughout an entire year. In reality, lengthy pedestrian counting periods are impractical because of the time and expense associated with counting (Soot, 1991; Davis et al., 1988; Cove and Clark, 1993; Hocherman et al., 1988).

Various methods of estimating pedestrian volume at a site have been developed in order to reduce the burden of data collection. Some of these strategies do not rely on direct sampling of pedestrian activity, and instead attempt to estimate the activity from land use variables, using similar techniques to the trip-generation methods used to predict vehicle travel (Hottenstein et al., 1997; Otis et al., 1995).

Other strategies rely on extrapolation procedures that convert short pedestrian counts into multi-hourly, daily, or annual estimates of pedestrian flow. There are two main strategies used to achieve the extrapolation of short pedestrian counts. One of these was used by Davis et al. (1988) in Washington, D.C. Pedestrian counts collected at 14 sites over three days were used to develop a set of equations relating

short count sample periods of 5, 10, 15 or 30-minutes to expansion periods of 1, 2, 3 and 4 hours. The equations were then validated using data from the remaining sites. It was found that the sample period should be in the middle of the period being sampled, and that the longer the sample period, the more accurate the estimate. The percent error in the estimate ranged from 11.9 percent to 33.6 percent depending on the length of the sampling period.

Although the procedure used by Davis et al. (1988) holds promise, it has some disadvantages. It does not take into account the time of day that the sample was taken, and does not differentiate between different types of sites. It also requires that samples be taken several times during the day in order to obtain a daily estimate.

The second procedure commonly used to extrapolate pedestrian counts involves the development of hourly conversion factors that can be used to expand any hour-long pedestrian count into a daily volume. Because this procedure is relatively simple vis a vis the method used by Davis et al. (1988), and because it takes into account the time of day and the characteristics of the site at which the sample was taken, it has been recommended as a means to extrapolate pedestrian volumes (Soot, 1991).

Moreover, the technique shares some characteristics with the methods of extrapolating short vehicle counts outlined in the Federal Highway Administration's Traffic Monitoring Guide (FHWA, 2001), which will be discussed below.

The remainder of this section focuses on the second method, which we refer to as the "factoring" method, although it has no specific name in the literature. The factor method involves tracking the temporal and spatial variations in pedestrian volumes in a given area and using them to expand a sample of short pedestrian counts into an annual measure of pedestrian volume

6.2. Temporal and Spatial Variations in Pedestrian Volumes

The factoring method of extrapolating pedestrian counts relies on knowledge or assumptions about how volumes fluctuate at the study site (Soot, 1991). This information is used to create hourly conversion factors that represent each hour's contribution to the daily flow. For example, if pedestrian flow at a site is perfectly constant, then each hour makes up 1/24, or 4.2 percent, of the day's total.

An hour-long count taken at any site could then be divided by .042 to obtain the daily total. The equation (4) shows the hourly adjustment factor in homogenous pedestrian flow (Zeeger et al., 2005). Similarly, if pedestrian volume were perfectly constant throughout the year, then a day long pedestrian count could be multiplied by 365 to obtain the yearly total.

$$\text{24 Hour Pedestrian Volume} = \text{Hour long count} / 0.042 \qquad (5)$$

The example of homogenous pedestrian flow is useful for illustrative purposes, but does not correspond to reality. Pedestrian volumes are known to fluctuate through time. The pedestrian volume distribution pattern at any given site varies from day to day according to diverse factors such as random variation in weather and day of the week (Hocherman et al., 1988; Hottenstein et al., 1997).

Cameron (1976) found that shopping areas in Seattle, Washington have higher levels of pedestrian activity during the dry summer months, the back-to-school season, and the holiday season, and lower levels during the rainier winter months. On the other hand, areas with little seasonal climate change have little seasonality in pedestrian volume (Hocherman et al., 1988).

In addition to these temporal fluctuations, there are also spatial variations in pedestrian volume. The daily pedestrian volume distribution pattern at one crosswalk may be different from that at a neighboring crosswalk, or in a crosswalk across town. Variations in the volume distribution through space may be produced by land uses surrounding the site (Davis et al., 1988) and the type of pedestrian activity associated with the site (Cameron, 1976).

Although each site is unique, some sites share similar patterns. The unique pattern at a site is sometimes called a "signature" (Soot, 1991). The most comprehensive review of pedestrian volume fluctuation patterns to date was undertaken in 1976 by Cameron. Several hundred days of data were collected, making it possible to track hourly, daily, and seasonal variations in pedestrian volume at each of the sites. It was found that the sites exhibited regular daily and hourly volume fluctuation patterns, and that similar types of sites tended to have similar volume distribution patterns (Cameron, 1976).

Similarities in the pedestrian volume distribution pattern at different sites can be exploited for the purpose of pedestrian volume estimation. Sites which are expected to share a similar pedestrian volume distribution can be treated as a group in order to facilitate the volume estimation process.

If the volume distribution for one site in the group is known, then it can be assumed that all sites in the group share the same distribution pattern. For example, Cameron (1976) classified pedestrian areas by the type of activity at the site: shopper, employee, visitor, mixed, commuter, and special, and identified characteristic pedestrian volume trends for each type of site. Zeeger et al. (2005) grouped sites on the basis of their location in a central business district, residential, or fringe area.

The following section describes how to apply the factoring method using a series of steps. The method involves grouping the sites in an area into strata that share similar pedestrian characteristics, making it similar to the stratified sampling techniques discussed in chapter 5.

6.3. Guide to Estimating Annual Pedestrian Exposure Using the Factoring Method

6.3.1. Select study area

Defining the target area for pedestrian volume monitoring is the first step in performing the factor analysis. Although the analysis can be performed at nearly any geographic scale, it is likely to be most feasible for jurisdictions such as large cities, metropolitan areas, Caltrans Districts, or the state. This is because the procedure requires all-day pedestrian counts, and the time and monetary investments required to collect this data may be harder to justify for small jurisdictions. Larger jurisdictions could achieve a statistical economy of scale by developing adjustment factors applicable to all areas (cities, counties, etc).

However, it is important to be aware of potential tradeoffs between the quality of the results and the size of the study area. One of the sources of error in the calculation of adjustment factors results from differences in the pedestrian volume fluctuation patterns within strata.

Large areas are more likely to contain heterogeneous pedestrian environments that will introduce error into the strata. For example, the city of San Francisco is characterized by mixed land uses, a grid-like street pattern, and high-density development. If one defined three strata within the city (e.g. residential area, employment area, and mixed), one would expect the pedestrian volume fluctuation patterns within these groups to be relatively homogenous, given the consistent character of the urban environment. However, if one defined the same three strata for the entire nine-county San Francisco Bay Area, one would expect a great deal more variation to occur within the strata, and therefore a great deal more error in the resulting volume estimate. Of course, larger jurisdictions may have the resources to account for these variations by selecting and sampling a larger number of strata.

6.3.2. Choose strata (employment center, residential area, mixed/fringe)

As described in the preceding literature review, areas in which the daily pedestrian volume fluctuation pattern is expected or assumed to be homogenous can be grouped into one or more strata. The raw pedestrian volumes at these sites may vary, but similarities in the surrounding land uses, intensity of development, and character of the pedestrian environment create similar temporal variations in pedestrian activity. The strata should be spatially defined, mutually exclusive, and should together equal the study area (Table 6.1). In other words, strata should be defined such that any site in the study area belongs to no more than one strata.

Table 6.1: Characteristics of Strata

Previous studies have grouped sites by the dominant land use, such as residential, central business district, and fringe area (Zeeger et al., 2005); or by the dominant type of pedestrian at the site, such as shopper, commuter, employee, visitor, and mixed (Cameron, 1976). The ideal selection of strata would account for all the possible sources of variation in activity, and would create a separate stratum for each pedestrian volume fluctuation pattern. The

1. Strata are defined by environmental or density variables
2. Sites within each strata are expected to have similar daily pedestrian volume fluctuation patterns
3. Strata are mutually exclusive
4. Sum of strata equal to entire study area

study of Davis et al. (1988), for example, found six unique pedestrian volume variation patterns among fourteen studied sites (Figure 6.1).

It is usually necessary to limit the number of strata groups selected, since each one requires a certain number of samples and is thus associated with a certain cost. This guide proposes three strata, though that number can be increased or decreased depending on the resources available, the desired accuracy of the estimate, and the heterogeneity of the study area.
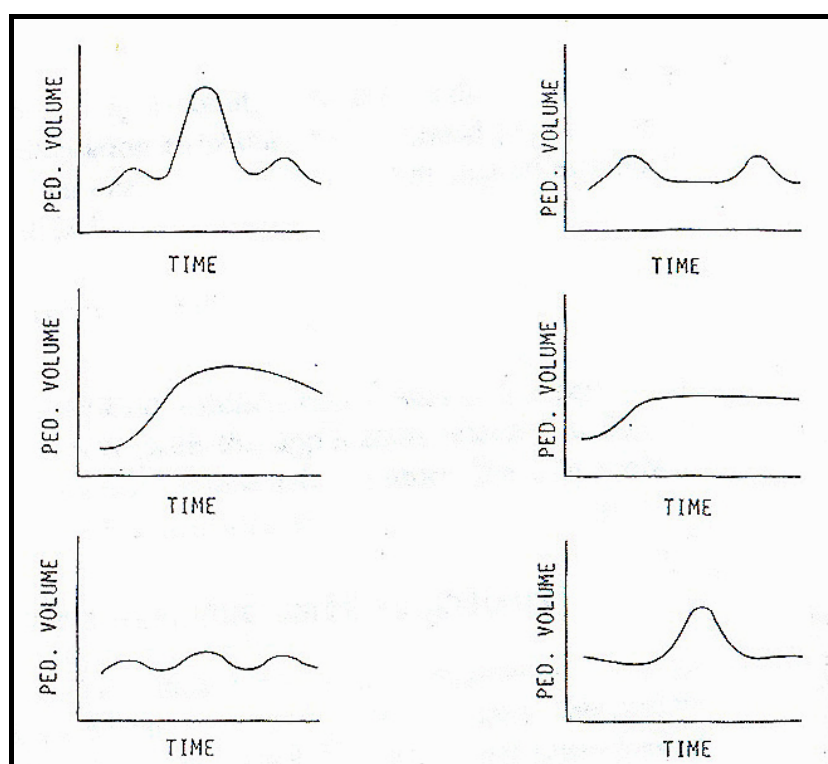
Figure 6.1: 12-hour Pedestrian Volume Distribution Patterns at Sites in Washington, D.C. (Davis et al., 1988)

Although the strata can be defined in a variety of ways, this guide proposes that they be defined in terms of their residential and employment density. The use of residential and employment density has three advantages. First, these data can serve as a simple proxy for more complicated measures of land use mix. Second, these data are readily available for through the U.S. Census Transportation Planning Package. Data may also be drawn from other local or national government sources, such as County Business Patterns data collected by the U.S. Census Bureau. Third, these data can be used to quantitatively define mutually exclusive strata. The

definition of strata by land-use type alone (e.g. CBD, residential) is more subjective and is not guaranteed to create groups that are mutually exclusive and sum to the entire study area.

The equation (5) may be used to assign areas to strata on the basis of area density. The formula and list of area types were developed by the Metropolitan Transportation Commission for use in regional transportation demand modeling (MTC, 1997). The original six categories of area type used by MTC are provided, as well as a simplified three-group area type that may be used for this study in Table 6.2.

$$AreaDensity = \frac{P + 2.5E}{AC + AI + AR} \qquad (6)$$

where,

P is total resident population within the target area

E is the total employment within the target area

AC is the commercial acreage within the target area

AI is the industrial acreage within the target area

AR is the residential acreage within the target area

Table 6.2: Categories of Area Type

Six-Group MTC Area Type	Simplified Three-Group Area Type
0 Core (Area Density > 300.0)	Central Business District (Area Density > 100.0)
1 Central Business District (Area Density = 100.0 - 300.0)	
2 Outlying Business District (Area Density = 55.0 - 100.0)	2 Fringe area (Area density = 30.0 – 100.0)
3 Urban (Area Density = 30.0 - 55.0)	3 Suburban and Rural (Area density = 6 – 30)
4 Suburban (Area Density = 6.0 - 30.0)	
5 Rural (Area Density < 6.0)	

6.3.3. Choose number of factors (hour, day, season, month, year)

The selection of strata described above reflects the need to account for *spatial* variation in pedestrian volume. This section describes the selection of adjustment factors which account for *temporal* variation in pedestrian activity.

The adjustment factors within a stratum will be used to develop an equation relating a given short count to an estimate of annual pedestrian volume for sites in that

stratum. The simplest equation for converting a short count (hourly) into an estimate of annual volume requires a single adjustment factor. This factor must reflect the proportion of the daily volume that the hour makes up in a specific stratum.

The number of adjustment factors required depends on the degree to which the pedestrian volume distribution pattern is expected to change throughout the hour, day, season, month, and year. That is, if the site is located in an area that has significant day-of-week or seasonal variations in pedestrian volume, additional adjustment factors may be necessary to account for those variations. For example, if the short count is taken in a cold month when pedestrian activity is diminished, then simply multiplying the daily estimate by 365 will result in an underestimate of pedestrian activity for the year. A seasonal adjustment factor would help correct for decreases in pedestrian volumes during winter months.

The extent of day-of-week and seasonal variation in pedestrian activity can be estimated by conducting all-day counts of pedestrian activity at a site on several days spread throughout the week and year. The results of such a study could be used to develop adjustment factors that could apply to all the strata, assuming that all strata are similarly affected by day-of-week and seasonal fluctuations in pedestrian volume.

The number of adjustment factors used also depends on resources. Increasing the number of adjustment factors will likely produce a better estimate of annual pedestrian volume, but will require additional sampling to implement. If limited resources make it impossible to develop day-of-week and seasonal adjustment factors, the study can be limited by collecting all counts during a specific time of year (e.g. early fall) and on a specific day (e.g. weekday or weekend day).

6.3.4. Calculate number of day-long counts needed

The number of day-long counts needed within each stratum is a function of the variability of the volume distribution within the stratum. To determine this, a pilot test should be conducted at a sample of sites throughout the study area.

It may occur that there is a great deal of variation in the data collected for each stratum. In this case, the definition of the strata should be examined and possibly

readjusted so that each stratum represents, as much as possible, sites within similar pedestrian volume distribution patterns. To facilitate this readjustment, detailed information should be collected on each site sampled during the day-long counts, including the surrounding land uses and type of pedestrian activity.

6.3.5. Collect day-long counts at sites

Day-long counts are collected at sites in the study area in order to determine the daily pedestrian volume fluctuation pattern at each site, which will reflect the daily pattern for all sites in the strata. In theory, it would be ideal to collect day-long counts on every day of the week for the year to determine daily, weekly, and seasonal volume fluctuation patterns for the strata. If this is not possible, then efforts should be made to be consistent in the day chosen for day-long counts. For example, it would be problematic to collect some day-long counts on Friday and others on Tuesday, as the volume distribution pattern will likely differ on each day of the week. Data collection should be avoided on anomalous days of the year, such as holidays, or during times of severe or uncharacteristic weather patterns.

In some cases, lack of automated counting equipment or sufficient resources may make it impractical to collect an entire 24-hour count of pedestrian volume. In these cases, it is advised that 15-hour counts be taken from 7:00 a.m. to 10:00 p.m. Hocherman et al. (1988) found that the period between 10:00 p.m. and 7:00 a.m. represents 3 percent of the daily volume in residential areas and 7 percent of the daily volume in the central business district.

The final result of the data collection should be a table indicating, for each stratum, the mean share of daily volume comprised by each hour in the day, as well as the standard deviation of the sample for each hour.

6.3.6. Develop factor equation

As noted above, the exact form of the factor equation depends on the number of adjustment factors developed during the sampling process. Assuming that only an hourly adjustment factor was developed, the factor equation would yield an average daily volume estimate for a specific day. The factor equation 6 would be used in this case. If a seasonal adjustment factor is developed, then equation 7 can be used (adapted from Hocherman et al., 1988).

$$Aadpv = Cij * K * Di \qquad (7)$$

where

AAdpv = Average daily pedestrian volume for site in strata f

Cij = short-count value in hour i and season j for site in strata f

Di = daily expansion factor for hour I in strata f

K = hourly multiplier: 60/minutes of short count (if less than a one-hour short count is taken)

$$AAdpv = Cij * K * Di * Sj \qquad (8)$$

where

AAdpv = Average daily pedestrian volume in strata f

Cij = short-count value in hour i and season j for site in strata f

K = hourly multiplier: 60/minutes of short count

Di = daily expansion factor for hour I in strata f

Sj = Seasonal correction factor for season j in strata f

6.3.7. Determine optimal length and time period of short count

Although the short count may be taken at any time of day, certain times of day may produce more accurate results. The chosen duration of the short count period will also influence the accuracy of the results and will affect the efficiency of the study.

Length of short count. The optimal length of the short count period is a function of the pedestrian volume at the site and the desired level of accuracy. Haynes (1977) found that the accuracy of a given counting period increases with the volume of pedestrians such that a shorter counting period is required at a high-volume site. A series of curves were developed to aid in the choice of counting period, as shown in Figure 6.2. The curve illustrates that, for example, an hour-long short count does not produce significantly less error than a 40-minute short count in areas with very high pedestrian volume (50 ped / minute).

These curves will be most helpful in urban areas with substantial numbers of pedestrians and will not apply in areas with low numbers of pedestrians. In these areas, it is possible that no pedestrians will be recorded within an hour-long period,

resulting in an erroneous average daily pedestrian volume of zero, even if the sample is taken during a peak travel period. To cope with this problem, three possible solutions are proposed:

✓ Collect more than an hour of pedestrian volume;

✓ Replace the count of zero with a count of .25. This method was used by Zeeger et al. (2005) at sites where an hour-long count produced zero pedestrians. It reflects the fact that pedestrian volume is very low without being zero.

✓ Use an alternative method. As noted above, several hours of data may be necessary to develop volume estimates at sites with few pedestrians. When counting pedestrians for several hours is impractical, an alternate method may be required, such as multiple regression techniques (Qin and Ivan, 2001).
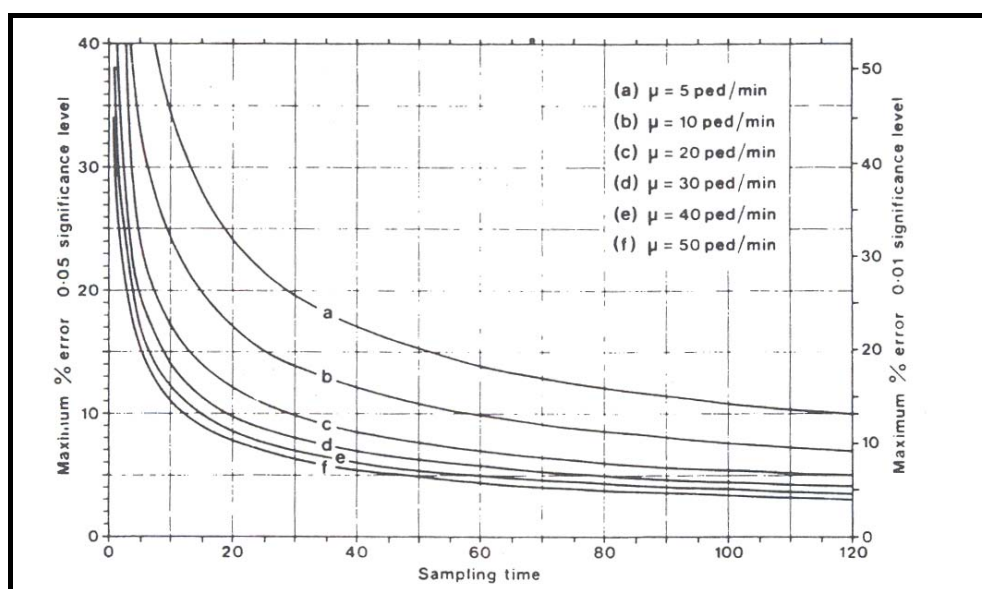
Figure 6.2: Relationship between maximum expected sampling error and sampling time for various levels of pedestrian activity (Haynes, 1977)

Time period of short-count. Three factors should guide the choice of when to sample the short-count at the study site:

✓ The expected or known peak hour of pedestrian volume at the site. As noted above, higher pedestrian volumes at a site may reduce the required length of the

short count and/or improve the accuracy of the short count. For that reason, there is a benefit to sampling pedestrian volume when volumes are expected to be at their highest.

✓ The standard deviation of the hourly adjustment factor. The hourly standard deviations developed for each hourly adjustment factor should be reviewed before sampling short counts. If one or more of the hours was shown to have a high standard deviation, efforts should be made to avoid sampling during that hour, as doing so will produce a less accurate result than sampling during an hour with a lower standard deviation.

✓ Sampling schedule. In order to economize resources available for the study, it is important to design a careful sampling schedule. The schedule should minimize the time lost to travel between sample sites. It is also possible to conserve additional time and resources by coordinating the pedestrian volume sampling schedule with vehicle volume sampling schedules (Schneider et al., 2005).

6.3.8. Calculate the error of the estimate

The accuracy of the estimation depends on several factors. Principal among these is the variability of pedestrian volumes at the site. Every real-world site is subject to some random day-to-day variation, but some sites are much more erratic than others. If the flow varies significantly, then a given count is less likely to be representative of the average flow.

Pedestrian volumes in residential areas in Israel were shown to have hourly standard deviation of 2 – 3.5 percent of the daily volume, whereas volumes in central business district were more stable, with a standard deviation of between 1 and 3.5 percent of the daily volume. In addition, pedestrian volumes taken during non-peak periods were shown to be more stable than those taken during peak periods (Hocherman et al., 1988). Thus the problem of random variation in pedestrian volume can be mitigated somewhat by collecting counts during time periods that tend to have less variation in pedestrian volumes, such as non-peak periods.

Error in the factored estimate is also generated by the process of grouping sites on the basis of expected, rather than empirically measured, similarity in the pedestrian

volume distribution patterns. Although sites with similar land uses may show similar pedestrian activity, there is likely to be great diversity within the grouping of "central business district", for example. This diversity introduces error into the volume estimate. The amount of error will depend on the extent of diversity within the group. Increasing the number of groups has the potential to decrease the error of volume estimates within each group.

Hocherman et al. (1988) summarized the sources of error in the factoring process with the following equation:

$$Var(aadpv) = K2 \ X \ f[var(Cij), \ var(Di), \ var(Sj) \qquad (9)$$

where:

Var(aadpv) is the variation in the average daily pedestrian volume

K^2 is the square of the hourly adjustment factor

Var(C) is the random day-to-day variation in any given hourly count

Var(D) is the deviation of the daily volume distribution at the location being studied from the volume distribution used to calculate the adjustment factor. It is a function of the homogeneity of sites within the strata

Var(S) is the variation of the seasonality factor used to correct for seasonal variations in pedestrian volume

Another source of error not included this equation is the error that occurs as adjustment factors become outdated. The adjustment factors developed for a group of sites may change from year to year as pedestrian distribution patterns are altered by changing land uses and pedestrian behavior. The extent of this error will depend on the frequency in which adjustment factors are recalculated.

6.3.9. Recalibrate equation

The power of the short-count expansion equation is derived from the assumption that pedestrian activity patterns remain relatively static over time. Over a period of years, however, pedestrian activity patterns will change in response to changing land uses and infrastructure. A site that was once primarily residential may be converted to office uses, for example, resulting in a surge of lunchtime pedestrian activity. Therefore, areas where new or infill development is occurring rapidly should

recalibrate more frequently (e.g. every 3 – 5 years) than areas with little development (e.g. 5 –10 years).

6.4. Example Expansion Procedures

This section provides two examples from the literature that used the factoring method described above to estimate pedestrian volumes.

6.4.1. Crosswalk study

The first example comes from a study of 2,000 uncontrolled crossings performed by Zeeger et al. (2005). The crossings were grouped into three types: sites in the central business district (CBD); sites in a fringe area; and sites in a residential area. Sites within each type were assumed to have similar daily pedestrian volume distributions.

Hourly adjustment factors were developed for the three types of sites through the collection of all-day (8- to 12- hour) counts at 22 of the 2,000 sites, as illustrated in Figure 6.3. Counts were not taken during the night time hours (7pm to 7am), but were estimated to represent about 14 percent of the daily total at the site. This estimation was based on the work of Cameron (1976) which found that that the period from 7pm to 7am comprises 14 percent of the 24-hour daily volume at a site. Similarly, Hocherman et al. (1988) found that this period makes up 14.9 percent of the daily volume in residential areas and 18.3 percent of the daily volume in CBD areas.

The pedestrian crossing volume at the remaining 2,000 sites was determined by multiplying a single hour-long count taken at the site by the hourly adjustment factor for that site. Then the daily volume was multiplied by 365 to obtain a yearly volume.

6.4.2. Study of pedestrian volumes in Israel

Hocherman et al. (1988) examined daily pedestrian volume distributions at 72 residential sites and 14 central business district sites in Haifa and Givatayim, Israel, to determine whether the factoring method could be used effectively to extrapolate short pedestrian counts.

It was found that the daily volume distributions at the residential sites were very similar and could be used to calculate an average daily pedestrian distribution at residential sites. The volume distributions at sites in the central business district also showed a clear pattern, with the main differences from residential sites being a smaller morning peak period and a lower hourly variation in pedestrian volume. The authors compared their results with similar distributions in Germany and Australia, and found similarities between the three distributions. Figure 6.4 shows the results of the comparison between the pedestrian volume distributions in these three countries.
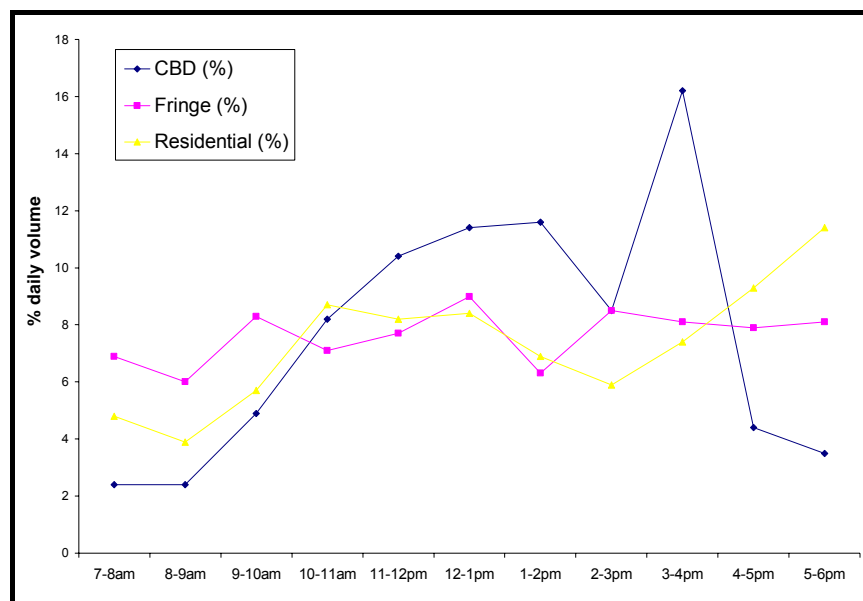
Figure 6.3: Daily volume adjustment factors developed for CBD, Fringe, and Residential Sites
(Zeeger et al., 2005)
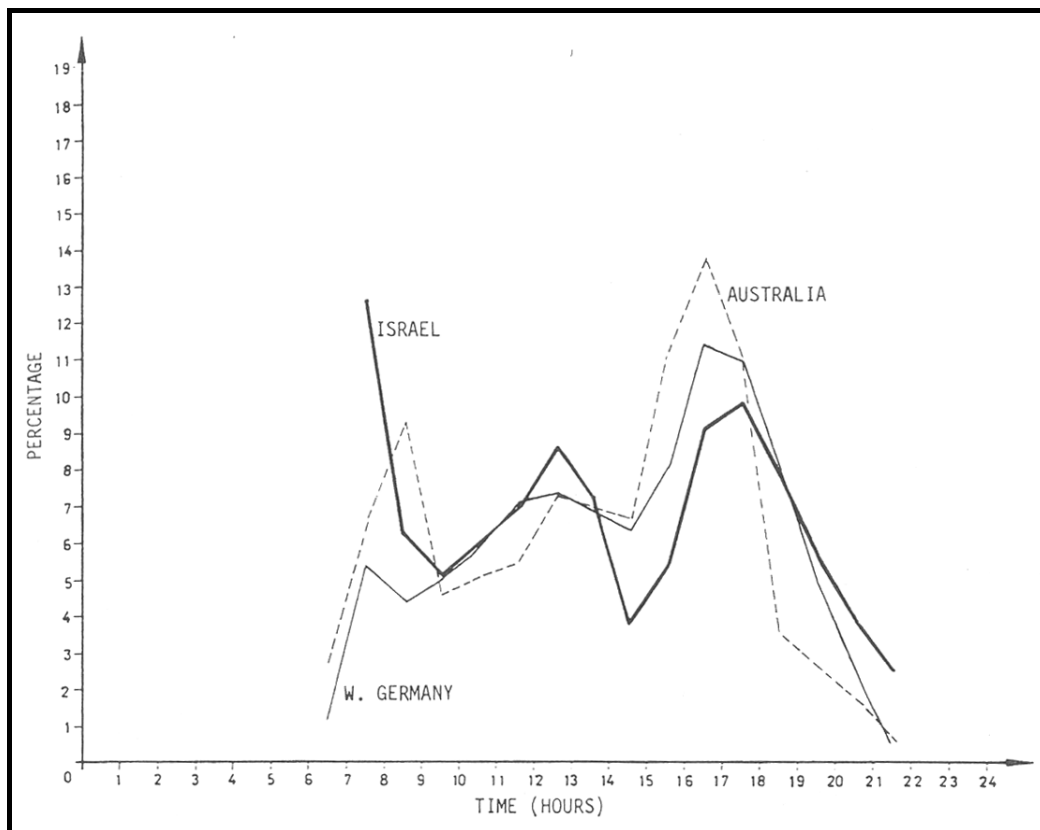
Figure 6.4: Comparison of daily pedestrian crossing volume distributions in Israel, Germany, and Australia (Hocherman et al., 1988)

6.5. FHWA Traffic Monitoring Guide

Although the volume monitoring procedures described in the Traffic Monitoring Guide (FHWA, 2001) involve vehicle volumes only, they employ the factoring method. The methods in the TMG are basically similar to the expansion methods described above, in that they rely on the development of factors to be applied to groups of similar roadways. However, the existence of readily available continuous counting devices makes the vehicle volume estimation process more statistically robust than the pedestrian volume procedures described above.

These devices, also known as Automatic Traffic Recorders (ATRs) are capable of recording volume fluctuation patterns continuously over a period of years. Pedestrian volumes, by contrast, are rarely collected for more than a period of hours or days at a time.

ATRs are typically placed in many locations throughout a state and are used in the development of time-of-day, day-of-week, and seasonal adjustment factors. The ATRs are then matched with groups of roadways on the basis of empirically measured similarities or by expected similarity on the basis of similar functional class or roadway type. The adjustment factors developed for a given group are used to convert short counts, usually of 48 hours or more, into measures of average annual daily traffic.

REFERENCES

Aggarwal, Y. P. **Better Sampling: Concepts, Techniques and Evaluation.** Sterling
Publishers Private Limited, 1988.

Alta Planning and Design. **National Bicycle and Pedestrian Document Project**.
Alta Planning and Design, Inc. www.altaplanning.com/. Accessed May 21,
2006.

Anderson, P.R., H.J. Montesin, and M.A. Adena. **Road Fatality Rates in Australia
1984-1985**. Federal Office of Road Safety, Canberra, Australia, 1989.

Banerjee, I. and D.R. Ragland. Evaluation of Countermeasures: A Study on the
Effect of Impactable Yield Signs Installed at Four Intersections in San
Francisco. In **TRB 86th Annual Meeting Compendium of Papers CD-ROM.**
Transportation Research Board, 2007.

Behavioral Risk Factor Surveillance System (BRFSS). National Center for Chronic
Disease Prevention and Health Promotion. **www.cdc.gov/brfss/about.htm.**
Accessed June 30, 2006.

Ben-Akiva, M. and S. Lerman. **Discrete Choice Analysis: Theory and Application
to Travel Demand.** The MIT Press, Cambridge, MA, 1985.

Bureau of Transportation Statistics (BTS) **Guide to Good Statistical Practice in the
Transportation Field.** U.S. Department of Transportation, Bureau of
Transportation Statistics, 2003.

Bureau of Transportation Statistics (BTS) **BTS Statistical Standards Manual.** U.S.
Department of Transportation, Bureau of Transportation Statistics, 2005.

California Department of Transportation (Caltrans). **California Statewide
Household Travel Survey**, **2000-2001.** Final Report. California Department
of Transportation, Division of Transportation System Information 2002.

California Department of Transportation. **California Motor Vehicle Stock, Travel and Fuel Forecast**. California Department of Transportation, Division of Transportation System Information, 2005.

California Highway Patrol. **Annual Report of Fatal and Injury Motor Vehicle Traffic Collisions.** 2004.

Cameron, R.M. Pedestrian Volume Characteristics. Institute of Transportation Engineers **Compendium of Technical Papers**, 1976.

Cameron, M.H. A Method of Measuring Exposure to Pedestrian Accident Risk. **Accident Analysis and Prevention**, Vol. 14, 1982, pp 397-405.

Campbell, B.J., C.V. Zegeer, H.H. Huang and M. J. Cynecki. **A Review of Pedestrian Safety Research in the United States and Abroad.** Publication FHWA-RD-03-042. Office of Safety Research and Development, Federal Highway Administration 2004.

Central London Partnership (CLP) **Automatic Pedestrian Counting Trial**. Stage 3 - Final Report. Central London Partnership, 2005.

Chan, C., F. Bu, and S. Shladover. **Experimental Vehicle Platform for Pedestrian Detection**. California PATH Research Report UCB-ITS-PRR-2006-16, 2006.

Chu, X. The Fatality Risk of Walking in America: A Time-Based Comparative Approach. **Walk21 Conference**: **Health, Equity and the Environment**, Portland, Oregon, 2003.

Chu, X. Measuring the Risk of Walking. Lecture. XVIIth Entretiens Jacques Cartier, EJC Conference, **Urban Road and Public Transit: Who must pay?**, Montreal, 2004.

Cove, L.L. and J.E. Clark. Estimating Pedestrian Volumes on College Campuses. **Transportation Research Record** 1405, Transportation Research Board, 1993, pp. 43-48.

Davis, S.E., L.E. King, and D.H. Robertson. Predicting Pedestrian Crosswalk Volumes. **Transportation Research Record** 1168, Transportation Research Board, 1988, pp 25-30.

Dharmaraju, R., D.A. Noyce, and J.D. Lehman. An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicycles. In **The 71st ITE Annual Meeting. Compendium of Technical Papers**. Institute of Transportation Engineering, Chicago, IL, 2001.

Diogenes, M.C., R.Greene-Roesel, L. Arnold, and D.R. Ragland. Pedestrians Counting Methods at Intersections: A Comparative Study. In **TRB 86th Annual Meeting Compendium of Papers CD-ROM.** Transportation Research Board, 2007.

European Transport Safety Council (ETSC). **Exposure Data for Travel Risk Assessment: Current Practice and Future Needs in the EU**. European Transport Safety Council, Brussels, Belgium, 1999.

Evans, L. **Traffic Safety and the Driver**. Van Nostrand Reinhold, New York, 1991.

Federal Highway Administration (FHWA). **Traffic Monitoring Guide**. Publication FHWA-PL-01-021. U.S. Department of Transportation, Federal Highway Administration, Office of Highway Policy Information, 2001.

Fehr and Peers Transportation Consultants. **City of Sacramento Pedestrian Master Plan: Making Sacramento the Walking Capital** (Public Review Draft) City of Sacramento: Sacramento, 2005.

Fitzpatrick, K., Turner, S. Brewer, M. Carlson, P., Ullman, B. Trout, N. Park, E.S., Whitacre, J., Lalani, N. and Lord, D. **Transit Cooperative Research Program Report 112/ National Cooperative Highway Research Program Report 562**, "Improving Pedestrian Safety at Unsignalized Crossings." Washington, DC: Transportation Research Board, 2006.

Hakamies-Blomqvist, L. Older Drivers' Accident Risk: Conceptual and Methodological Issues. **Accident Analysis and Prevention**, Vol. 30, 1998, pp 293-297.

Hakkert, A.S., and L.Braimaister. **The Uses of Exposure and Risk in Road Safety Studies**. Publication R-2002-12. SWOV Institute for Road Safety Research.. Leidschendam, the Netherlands, 2002

Hauer, E. Traffic Conflicts and Exposure. **Accident Analysis and Prevention**, Vol. 14, No 5, 1982, pp. 359-364

Hauer, E. On Exposure and Accident Rate. **Traffic Engineering and Control**, Vol 36, No. 3, 1995, pp. 134-138.

Hauer, E. Computing and Interpreting Accident Rates for Vehicle Types or Driver Groups. **Transportation Research Record**, Vol. 1746, 2001, pp 69-73.

Haynes, R. Sampling Time and Sampling Error in Pedestrian Counts. **Traffic Engineering Control**, Vol. 18, No. 2, 1997, pp. 72-74.

Herms, B.F. **Pedestrian Crosswalk Study: Accidents in Painted and Unpainted Crosswalks**. Police Department, Traffic Bureau, San Diego, 1970.

Hocherman, I., A.S. Hakkert and J. Bar-Ziv. Estimating the Daily Volume of Crossing Pedestrians from Short-Counts. **Transportation Research Record** 1168, Transportation Research Board, 1988, pp. 31-38.

Høj, N.P. and W. Kröger. Risk Analyses of Transportation on Road and Railway from a European Perspective. **Safety Science**, Vol. 40, No. 1-4, 2002, pp. 337-357.

Hottenstein, A., S. Turner and G. Shunk. **Bicycle and Pedestrian Travel Demand Forecasting: Summary of Data Collection Activities.** Publication FHWA/TX-98/1723-1. Texas Transportation Institute, Federal Highway Administration, 1997.

Hunter, W., J. Stutts, W. Pein and C. Cox. **Pedestrian and Bicycle Crash Types of the Early 1990s**. Publication FHWA-RD-95-163, U.S Department of Transportation, Federal Highway Administration Office of Traffic Safety and Operations, 1996.

Ivan, J. N., P. J. Ossenbruggen, X. Qin and Jyothi Pendarkar. **Rural Pedestrian Crash Rate: Alternative Measures of Exposure**. Project UCNR 11-10. United State Department of Transportation. Region I University Transportation Center, 2000.

Jacobsen, P. L. Safety in Numbers: More Walkers and Bicyclists, Safer Walking and Bicycling. **Injury Prevention.** Vol 9, 2003, pp 205-209.

Jonah, B.A. and G.R. Engel. Measuring the Relative Risk of Pedestrian Accidents. **Accident Analysis and Prevention**, Vol. 5, No. 3, 1983, pp.193-206.

Jorgensen, N.O. The Risk of Injury and Accident by Different Travel Modes. **Proceedings of the International Conference on Passenger Safety in European Public Transport**. Transport Safety Council, Brussels, 1996, pp. 17-25.

Keall, M.D. Pedestrian Exposure to Risk of Road Accident in New Zealand. **Accident Analysis and Prevention**, Vol. 27, No. 3, 1995, pp. 729-740

Knoblauch, R.L., H.N. Tobey, and E.M. Shunaman. Pedestrian Characteristics and Exposure Measures. **Transportation Research Record**, Vol. 959, 1984, pp 35-41.

Knoblauch, R. L., M. T. Pietrucha, and M. Nitzburg. Field Studies of Pedestrian Walking Speed and Start-Up Time. **Transportation Research Record,** No. 1538, Pedestrian and Bicycle Research, 1996, pp 27-38

Komanoff, C., and C. Roelofs. **The Environmental Benefits of Bicycling and Walking. National Bicycling and Walking Case Study No. 15.** Publication FHWA-PD-93-015. FHWA, U.S. Department of Transportation, 1993.

Kraus, J.F., E.G. Hooten, K.A. Brown, C. Peek-Asa, C. Heye and D.L. McArthur. Child Pedestrian and Bicyclist Injuries: Results of Community Surveillance and a Case-Control Study. **Injury Prevention**, Vol. 2, 1996, pp. 212-218.

Lee, C. and M. Abdel-Aty. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. **Accident Analysis and Prevention**, Vol 37, 2005, pp. 775-786.

McNally, M.G. The Four-Step Model. In D.A. Hensher and K.J. Button, Ed. **Handbook of Transport Modelling. Pergamon**, Oxford, United Kingdom, 2000, pp. 35-52.

Metropolitan Transportation Commission (MTC) **Travel Demand Models for the San Francisco Bay Area (BAYCAST-90)**. Technical Summary. Metropolitan Transportation Commission, 1997

Midwest Regional University Transportation Center (MRUTC) **Development of a Guide to Statistics for Maintenance Quality Assurance Programs in Transportation**. Research Progress Report for the Quarter Ending: June 30, 2005.

Mitman, M. and D.R. Ragland. Toward Practice-Wide Standards for Pedestrian Data Collection: An Evaluation of Pedestrian Behavioral Data Collection Methods for Multi-Lane Roads with Low Pedestrian Volumes In **TRB 86th Annual Meeting Compendium of Papers CD-ROM.** Transportation Research Board, 2007.

Nassirpour, M. Overview of Sample Surveys (Presentation). Applied Sampling Workshop at the **Annual TRB Conference**, Washington, D.C. , 2004.

National Center for Statistics and Analysis (NHTSA). **Traffic Safety Facts: 2004 Data**. National Center for Statistics and Analysis, Washington, D.C., 2004.

National Household Travel Survey (NHTS). Bureau of Transportation Statistics, U.S. Department of Transportation. **nhts.ornl.gov/2001/index.shtml**. 2001 Accessed May 30, 2006.

National Household Travel Survey (NHTS). Bureau of Transportation Statistics, U.S. Department of Transportation. **[Website] http://nhts.ornl.gov/.** Accessed December 17, 2006.

Noyce, D. and R. Dharmaraju. **An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicyclists.** Massachusetts Highway Department report, 2002.

Noyce, D., A. Gajendranand and R. Dharmaraju. Development of a Bicycle and Pedestrian Detection and Classification Algorithm for Active-Infrared Overhead Vehicle Imaging Sensors. In **Transportation Research Board 85th Annual Meeting**. CD-ROM. Transportation Research Board, Washington , D.C., 2006

Ogden, K.W. The Effects of Paved Shoulders on Accidents on Rural Highways. **Accident Analysis and Prevention,** Vol. 29, 1997, pp. 353-362.

Ossenbruggen, P.J. The Impacts of a Safety Compliance in Highway Design. **Risk: Health, Safety and Environment**, Vol. 10, 1999, pp. 359–379.

Otis, S. C., R. B.Machemehl and H. S. Mahmassani. Predicting Pedestrian Volumes Based on Land Use: a Methodology. **Fifth National Conference on Transportation Planning Methods Applications**. A Compendium of Papers, Vol. 1. Transportation Research Board, 1995.

Pucher, J. and J.L. Renne. Socioeconomics of Urban Travel: Evidence from the 200NHTS. **Transportation Quarterly**, Vol. 7, No. 3, Summer 2003, pp. 49-77.

Pucher, John, and Lewis Dijkstra. Making Walking and Cycling Safer: Lessons from Europe.**Transportation Quarterly**, Vol. 54, 2000, pp 25-50.

Pucher, J. and L. Dijkstra Promoting Safe Walking and Cycling to Improve Public Health: Lessons from the Netherlands and Germany. **American Journal of Public Health**, Vol. 93, 2003, pp 1509-1516.

Qin, X. and J.N. Ivan. Estimating Pedestrian Exposure Prediction Model in Rural Areas. **Transportation Research Record 1773**, Transportation Research Board, 2001, pp 89-96.

Raford, N. and D.R. Ragland. Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: Case of Boston, Massachusetts. In **Transportation Research Board 85th Annual Meeting**. CD-ROM. Transportation Research Board, Washington , D.C., 2006

Raford, N. and D.R. Ragland. Space Syntax: Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety. In **Transportation Research Record: Journal of the Transportation Research Board**, No. 1878, TRB, National Research Council, Washington, D.C., 2004, pp. 66-74.

Roberts, I, R. Norton and B. Taura. Child Pedestrian Exposure Rates: The Importance of Exposure to Risk Relating to Socioeconomic and Ethnic Differences in Auckland, New Zealand. **Journal of Epidemiology and Community Health**, Vol. 50, 1996, pp 162-165.

Schneider, R., R. Patton, J. Toole and C. Raborn. **Pedestrian and Bicycle Data Collection in United States Communities: Quantifying Use, Surveying Users, and Documenting Facility Extent**. FHWA, Office of Natural and Human Environment, U.S. Department of Transportation, Washington, D.C., 2005.

Schwartz, W. and C. Porter. **Bicycle and Pedestrian Data: Sources, Needs, and Gaps.** Publication BTS00-02. Bureau of Transportation Statistics. U.S. Department of Transportation, Washington, D.C., 2000.

Schwartz, W.L., C. D. Porter, G. C. Payne, J. H. Suhrbier, P. C. Moe and W. L. Wilkinson III. **Guidebook on Methods to Estimate Non-Motorized Travel: Overview of Methods**. Publication FHWA-RD-98-165. FHWA, U.S. Department of Transportation ,1999.

Schweizer, T. Methods for counting pedestrians. In T**he 6th International Conference on Walking in the 21st Century**. Walk21-VI "Everyday Walking Culture". Zurich, Switzerland, 2005.

Senevarante, P. and J. Morall. Analysis of Factors Affecting the Choice of Route of Pedestrians. **Transportation Planning and Technology**, Vol. 10, No. 1, 1986, pp. 147-159.

Shapiro, G., D. Utter, W. Mix, J. Croos, D. Marker, and N. Bondy. Sample Design Issues for Surveys Involving the Observation of Drivers. **Proceedings of the Annual Meeting of the American Statistical Association**, 2001.

Som, R. K. **Pratical Sampling Techniques**. Second Edition, Revised and Expanded. Marcel Dekker, Inc, New York. 635 p. 1996.Som (1996)

Soot, S. Trends in Downtown Pedestrian Traffic and Methods of Estimating Daily Volumes. **Transportation Research Record 1325**, Transportation Research Board, 1991, pp 75-82.

Statistics Canada. **http://www.statcan.ca/start.html** Accessed November 10, 2006.

STPP. **Mean Streets 2004: How Far Have We Come?** Surface Transportation Policy Project, Washington, D.C., 2004.

 STPP. **Mean Streets**. Surface Transportation Policy Project, Washington, D.C. 2002.

Transport Research Laboratory (TRL). **Cost-Benefit Analysis of Measures for Vulnerable Road Users**. Deliverable 5 for European Research Project PROMISING. SWOV Institute for Road Safety Research, The Netherlands, 2001.

Transportation Research Board's (TRB) Committee on Urban Transportation Data and Information Systems (ABJ30). **Household Travel Surveys on the WWW** trb.mtc.ca.gov/urban/hhsurvey.html. Accessed July 17, 2006.

United States Census Bureau, **American Community Survey.** 2004. Tables generated using American FactFinder. factfinder.census.gov. Accessed July 17,2006.

United States Census Bureau. **American Community Survey 2005 Questionnaire.** 2005. http://www.census.gov/acs/www/Downloads/SQuest05.pdf. Accessed May 30, 2006.

United States Census Bureau. **American Community Survey Basics.** 2006 www.census.gov/acs/www/SBasics/index.htm. Accessed May 30, 2006.

Zeeger, C.V., K.S. Opiela, and J.J. Cynecki. Effect of Pedestrian Signals and Signal Timing on Pedestrian Accidents. **Transportation Research Record** 847, TRB National Research Council, Washington, DC, 1982, pp. 62-72

Zeeger, C.V., R. Stewart, H. Huang, P.A. Lagerwey, J. Feaganes and B.J. Campbell. **Safety Effects of Marked versus Unmarked Crosswalks at Uncontrolled Locations: Final Report and Recommended Guidelines**. Publication FHWA–HRT–04–100. Office of Safety Research and Development, Federal Highway Administration, 2005.

APPENDIX A: Example of a Tally Sheet Used to Count Pedestrian

Intersection: _____

Data Collected by: _____

Data Collected on: _____

Period: ( ) 1:00 to 1:30 pm ( ) 1:31 to 2:00 pm ( ) 2:01 to 2:30 pm ( ) 2:31 to 3:00 pm

( ) 3:01 to 3:30 pm ( ) 4:00 to 4:30 pm ( ) 4:31 to 5:00 pm

LEGEND:

Age options:	Direction options:	Gender options:
1 - =<12 2 - 13-18 3 - 19-25 4 - 26-35 5 - 36-50 6 - 51-64 7 - 65+	1 2	M - Male F - Female

PED #	DIRECTION		AGE	GENDER	PED #	DIRECTION		AGE	GENDER	PED #	DIRECTION	AGE	GENDER
1	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	6	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	1			
										2			
										3			
										4			
										5			
										6			
										7			
2	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	7	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	8			
										9			
										10			
										11			
										12			
										13			
										14			
3	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	8	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	15			
										16			
										17			
										18			
										19			
										20			
										21			
4	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	9	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	22			
										23			
										24			
										25			
										26			
										27			
										28			
5	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	10	( ) 1 ( ) 2		( ) =<12 ( ) 13-18 ( ) 19-25 ( ) 26-35 ( ) 36-50 ( ) 51-64 ( ) 65+	( ) Male ( ) Female	29			
										30			
										31			
										32			
										33			
										34			
										35			

Pedestrian Counting Methods at Intersections: a Comparative Study

Submission Date: August 1, 2006

Word Count: 2976

Number of Figures and Tables: 6

Authors:
Mara Chagas Diogenes
Traffic Safety Center
University of California, Berkeley
140 Warren Hall #7360
Berkeley, CA 94709
510-643-7625
maracd@berkeley.edu

Ryan Greene-Roesel
Traffic Safety Center
University of California, Berkeley
140 Warren Hall #7360
Berkeley, CA 94709
510-643-7625
ryangr@berkeley.edu

Lindsay S. Arnold
Traffic Safety Center
School of Public Health
University of California, Berkeley
140 Warren Hall
Berkeley, CA 94720
510-282-5896
larnold@berkeley.edu

David R. Ragland (corresponding author)
Traffic Safety Center
University of California, Berkeley
140 Warren Hall #7360
Berkeley, CA 94709
510-642-0655
510-643-9922 (Fax)
davidr@berkeley.edu

ABSTRACT

Resources for implementing countermeasures to reduce pedestrian collisions in urban centers are usually allocated on the basis of need, which is determined by risk studies. They commonly rely on pedestrian volumes at intersections. The methods used to estimate pedestrian volumes include direct counts and surveys, but few studies have addressed the accuracy of these methods. This paper investigates the accuracy of three common counting methods: manual counts using sheets, manual counts using clickers, and manual counts using video cameras. The counts took place in San Francisco. For the analysis, the video image counts, with recordings made at the same time as the clicker and sheet counts, were assumed to represent actual pedestrian volume. The results indicate that manual counts with either sheets or clickers systematically underestimated pedestrian volumes. The error rates range from 8-25%. Additionally, the error rate was greater at the beginning and end of the observation period, possibly resulting from the observer's lack of familiarity with the tasks or fatigue.

INTRODUCTION

Road collisions are a major public health concern throughout the world. It is estimated that 1.2 million traffic fatalities occur each year worldwide. The problem is especially acute for pedestrians, who face a significantly greater risk of death when involved in traffic collisions than do vehicle occupants *(1)*. Significant resources are focused on countermeasures that aim to reduce the risk of pedestrian injury. Because resources are limited, risk analysis is necessary to develop cost-effective countermeasures *(2)*.

Risk is defined as the frequency of an undesired event or collision per unit of exposure. Pedestrian volume is the exposure measure most frequently used in risk analysis. According to Gårder *(3)*, pedestrian risk should be calculated as a function of pedestrian volume, not just vehicle volume. Although many state, regional, and local agencies have developed methodologies to collect pedestrian volume data, there is no consensus on which method is best *(4, 5)*. To improve the risk monitoring process, it is necessary to define a systematic pedestrian counting method.

The two most frequent types of pedestrian counting methods are direct counts and surveys. Direct counts involve direct observation of pedestrian activity at fixed locations, such as crosswalks or intersections. Surveys indirectly capture pedestrian activity in a geographic area by gathering travel data from a sample *(6)*.

Pedestrian volumes at intersections are usually collected directly using either (i) manual counts, taken by collectors in the field, or (ii) automated counts using specialized equipment. Although motorized vehicles are commonly counted with automated devices, the technology for counting non-motorized modes of transportation, especially pedestrians, is not very developed *(7)*.

The accuracy of these counting methods directly affects the accuracy of the exposure estimate and thus the value of the risk analysis at an intersection. However, few studies have attempted to compare the accuracy of different counting methods. This paper aims to compare the accuracy of three common pedestrian counting methods: (i) manual counts using sheets; (ii) manual counts using clickers; and (iii) manual counts using video cameras.

METHODS

The research was conducted at 10 different intersections in the city of San Francisco, California, during the last two weeks of April and the first week of May, 2006. Field observers collected pedestrian counts with either sheets or manual clickers. Counts were taken for four hours between 1:00 pm and 6:00 pm, with a break of one hour. Video footage of the intersection was recorded simultaneously with the field counts.

Two persons were contracted from a private consulting firm specializing in data collection. One individual made the field observations, and the other operated the video recorder. The contracted staff was the same for all data collection. Sheets were used at eight intersections and clickers at two intersections. The selected intersections had different pedestrian flows, with values varying between 12 and 262 pedestrian crossings per hour based on the video analyses, as shown in Table 1. Figures 1 and 2 present the camera angles used at two of the study intersections.

TABLE 1 Data Collection Schedule and Pedestrian Flow

Intersection	Date	Method	Volume (ped)	Period (hours)	Flow (ped/hour)
France and Mission St.	04/17/2006	Manual with sheets	128	4	32
Admiral Ave. and Mission St.	04/18/2006	Manual with sheets	49	4	12
16th St. and Capp	04/19/2006	Manual with sheets	412	4	103
Geneva and Mission St.	04/20/2006	Manual with sheets	1046	4	262
Folson and 7th St.	04/21/2006	Manual with sheets	334	4	84
Harrison and 7th St.	04/24/2006	Manual with sheets	651	4	163
Market and Castro	04/25/2006	Manual with sheets	579	4	145
Market and Noe	04/26/2006	Manual with sheets	994	4	249
Harrison and 10th St.	05/03/2006	Manual with clickers	161	4	40
Santa Rosa and Mission St.	05/05/2006	Manual with clickers	338	4	85

Before the start of data collection, the researchers supplied the field staff with the following directions:

1. The data collection must be synchronized with the video. The person collecting the data should begin to count the pedestrians when the video begins to run. During the period that the tape is being changed, the observer should stop counting.
2. The field observer must note any problem or interruption in the data collection, such as a break or lack of attention for any reason. These interruptions are important since the main objective was to compare the accuracy of the methods.

3. The field observer must count only pedestrians who cross the street centerline (e.g. the middle of the crossing). He or she should not count bicyclists unless they are walking their bicycle across the intersection.
4. The field observer must stand close to the crosswalk.

Field data were entered into a Microsoft Access 2000 database. For quality control, all database tables were compared with the original field data sheets.

FIGURE 1 Camera angle used at Admiral Ave. and Mission St.

FIGURE 2 Camera angle used at Market and Castro (still from video tape)

Manual with sheets
The field observer received a sheet with three fields: (i) direction of travel; (ii) pedestrian gender; and (iii) age. The observer was instructed to use his best judgment to assign the pedestrian to one of seven age categories.

At the top of the sheet, the observer was instructed to write the following information: (i) name of the intersection; (ii) his/her name; (iii) date of the data collection; and (iv) period of the data collection (check box) – divided in periods of 30 minutes. The field observer was told to concentrate on accurately counting the number of pedestrians, even if it meant leaving gender and age fields blank in crowded intersections.

To improve the analysis, after the fourth day (April 20), the field observer was asked, when possible, to take note of any distinguishing characteristics that would allow an individual to be identified in the video, i.e., clothing color, hair color, parcels or suitcases, exact time, and so on. This information made it possible to determine when the field observer missed or over-counted pedestrians, and to determine whether the manual data collection was properly synchronized with the video.

Manual with clicker
On May 3 and May 5, the field staff collected pedestrian counts using a manual clicker. The observer clicked once for every pedestrian crossing the intersection, regardless of direction. At the end of every 10-minute period, the observer noted the count on the clicker on the data sheet provided.

Manual with Video
The intersections were videotaped using a camera set up on a flatbed truck parked opposite the crosswalk being studied. The camera recorded an image of the crosswalk at an angle that allowed both directions of pedestrian travel to be captured. Video tapes were replaced after each hour.

Researchers involved in the study carefully analyzed the video tapes in order to obtain the most reliable results possible. The researchers tried to identify each pedestrian counted by the field observer. This task was only possible for the days that the field observer noted individual pedestrian characteristics.

The tapes were viewed in variable time, and sometimes viewed more than once if the results were in doubt. On average, one hour of video tape required three hours of

video analysis. During the analysis, the researchers paid attention to whether the field counts were synchronized with the videotape and looked for any discrepancies between the field observations and the video images.

DATA ANALYSIS

The purpose of the data analysis was to compare the accuracy of the methods. Because it was not possible to know the exact number of pedestrians on the roadway at any given time, inter-reliability between the methods was used as a proxy for accuracy. The counts derived from the video tapes were assumed to be closest to the actual pedestrian volume.

The comparison used the relative difference between the counts taken through each method to calculate the error:

$$Error = \frac{NPi - NPv}{NPv} \quad (1)$$

where NP_i is the number of pedestrians counted in the field and NP_v is the number of pedestrians counted using the video images. The error was calculated for each interval of data collection (30 minutes for the sheets and 10 minutes for the clickers), as well as for the total number of pedestrians counted at each intersection.

Synchronization of the field counts and video taping was a major issue identified during the video analysis, despite the fact that field staff were directed to synchronize the counting methods. Sometimes the field observer began counting slightly before or after the video camera began recording. When this occurred, it was difficult to compare the counts obtained through each method. To improve the results of the comparison study, counts taken in periods when the field observer was not synchronized with the video were not included in the calculation of the intersection error.

Comparisons of the accuracy of pedestrian gender and age identification were also made, but not included in this paper. The researchers concluded that it was not possible to precisely identify the gender or age of the pedestrians from the video images because of low image resolution.

RESULTS

In the first week of data collection, the field observer did not follow all of the instructions he was given and did not consistently collect data for four-hour periods. For example, he sometimes started counting late; failed to take note of his breaks; and counted bicycles as pedestrians. Despite this, the video tapes were analyzed for the entire counting period (four hours) in order to determine the average hourly pedestrian volume (Table 1).

The results of the comparison reveal that the field observer systematically counted fewer pedestrians than were observed on the video recordings. The average error calculated for the manual counting using sheets was 15%, varying from 9% to 25%, as shown in Tables 2. For the manual counting with clickers, the average error was 11%, varying from 8% to 15% (Table 3). Given the variation in the results, it is not possible to determine which method, with sheets or clickers, is the most accurate.

TABLE 2 Comparison of Counting Methods (Video vs. Sheets)

Period	Date						
	4/17/2006	4/18/2006	4/19/2006	4/20/2006	4/21/2006	4/24/2006	4/25/2006
	Error	Error	Error	Error	Error	Error	Error
1:00 to 1:30	Not Counted	Not Counted	-27%	-28%	-16%	-7%	-22%
1:30 to 2:00	150%*	Not Counted	-18%	-6%	0%	-2%	-17%
2:00 to 2:30	-13%	0%	3%	-23%	-17%	-16%**	-29%
2:30 to 3:00	-14%	0%	-28%	-2%	-12%		-26%
4:00 to 4:30	-13%	-22%	-42%	-14%	-8%	-8%	-27%
4:30 to 5:00	-21%	86%*	-67%	-15%	-10%	-11%	-17%
5:00 to 5:30	Not Counted	Not Counted	-25%	-16%	-5%	-3%	-25%
5:30 to 6:00	Not Counted	Not Counted	-49%	3%	-8%	-10%	-31%
Error (Total)	-15%	-11%	-21%	-12%	-10%	-9%	-25%

* Not included in the total, because it was not synchronized with the video

**In this period, the field observer failed to record the counts in half hour periods

TABLE 3 Comparison of Counting Methods (Video vs. Clickers)

	5/3/2006						5/5/2006					
	1:00 to 2:00pm						1:00 to 2:00pm					
Error (10 min)	-11%	-43%	-13%	0%	0%	0%	0%	0%	-19%	17%	-8%	100%
Error (hour)	-11%						2%					
	2:00 to 3:00pm						2:00 to 3:00pm					
Error (10 min)	-25%	-67%	0%	100%	-50%	0%	0%	-14%	25%	-31%	-8%	9%
Error (hour)	-23%						-5%					
	4:00 to 5:00pm						4:00 to 5:00pm					
Error (10 min)	0%	17%	33%	-25%	-11%	0%	50%	-25%	-41%	-33%	-40%	-88%
Error (hour)	0%						-32%					
	5:00 to 6:00pm						5:00 to 6:00pm					
Error (10 min)	-20%	0%	38%	-33%	0%	20%	-30%	6%	-64%	-15%	-8%	-88%
Error (hour)	0%						-21%					
Error (4 hours)	**-8%**						**-15%**					

An in-depth analysis of the data revealed that error was often greater at the beginning and end of the data collection period. Possible explanations for this finding include: (i) the observer's lack of familiarity with the intersection and the counting method at the beginning of the data collection; (ii) the long counting periods, which may have caused the observer to become fatigued and lose attention; and (iii) lack of synchronization with the video that was not possible to identify.

It was assumed that the observer would have more difficulty counting at intersections with high volumes of pedestrians, increasing the error value. However the results revealed that pedestrian flow did not influence the error, since the correlation (R^2 =0.1) between them was weak. Figure 3 presents a graph with the relationship between the error and the pedestrian flow.
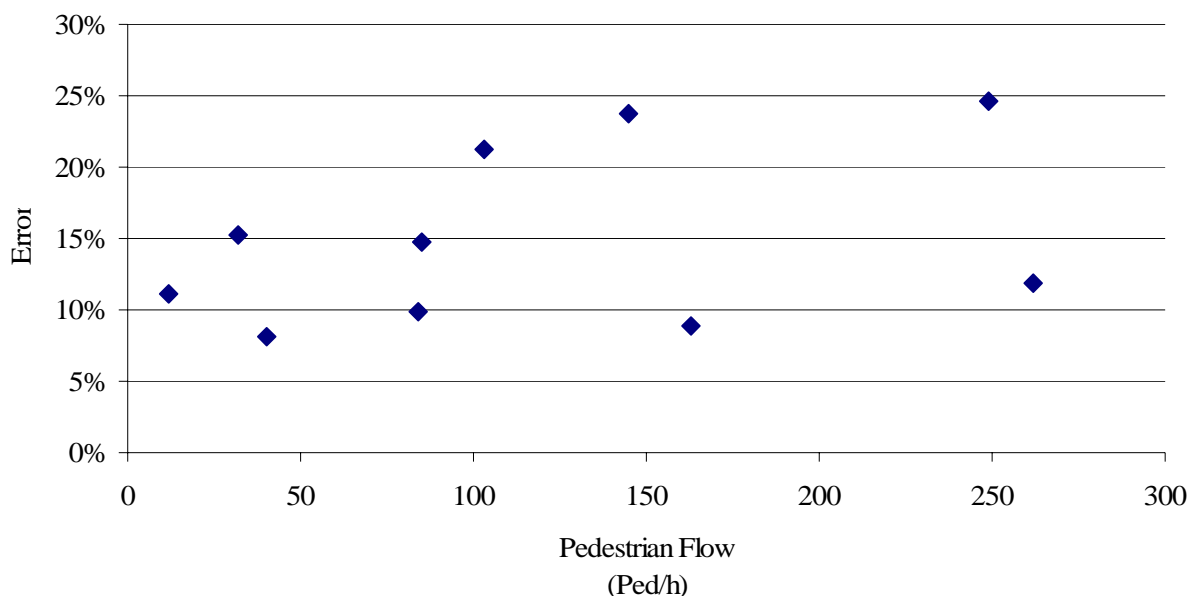
FIGURE 3 Relationship between the error and the pedestrian flow

DISCUSSION

The most significant results of this study were that pedestrian counts taken in the field were systematically lower than counts taken by observing video recordings, and that the accuracy of field counts did not seem to be strongly related to pedestrian flow. These results stem from the fact that the collection of field counts using either sheets or clickers is very difficult to control, and requires planning and organization during the counting day *(5)*.

The level of observer attention is one aspect of field data collection that is difficult to control. In this study, the observer may have become distracted at intersections with little pedestrian activity, but may have been more focused in areas with high activity that demanded his attention. It is also possible that the error was related to the observer's unique characteristics and motivation. Future studies should use multiple field observers to determine how the characteristics of the observers, such as their experience and background, affect the quality of the pedestrian counts. However, given the budgetary constraints of most transportation agencies, it may be difficult to ensure that field observers have high-level training and experience.

It was expected that manual counts taken with clickers would have very low error because this method allows the observer to keep his attention on the intersection

and does not demand that he identify and record pedestrian characteristics. No significant difference was found in the relative accuracy of manual counts using clickers and manual counts using sheets; however, more research is needed to compare the methods.

Although this study suggests that field counts may be less accurate than counts taken with video images, it is often necessary to use field observers to record detailed pedestrian characteristics and behaviors. It is difficult to identify these characteristics on video recordings without adequate image resolution and a well-selected camera angle.

This study suggests that video recordings should be used in situations where the accuracy of the count is of primary importance. However, users of this method should be aware that obtaining an accurate count from video can be very time consuming and requires meticulous attention to the video analysis. Overall, the choice of pedestrian counting method depends on the data collection needs and available resources.

ACKNOWLEDGEMENTS

REFERENCES
1. Peden, M., R. Scurfield, D. Sleet, D. Mohan, A.A. Hyder, E. Jarawan and C. Mathers. World Report on Road Traffic Injury Prevention. World Health Organization, Geneva, 2004.
2. Høj, N.P. and W. Kröger. Risk Analyses of Transportation on Road and Railway from a European Perspective. *Safety Science*, Vol. 40, No. 1-4, 2002, pp. 337-357.
3. Gårder, P.E. The impact of speed and other variables on pedestrian safety in Maine. *Accident Analysis & Prevention*, Vol. 36, No. 4, 2004, pp. 533-542.
4. Schneider, R., R. Patton, J. Toole and C. Raborn. P*edestrian and Bicycle Data Collection in United States Communities: Quantifying Use, Surveying Users, and Documenting Facility Extent.* FHWA, Office of Natural and Human Environment, U.S. Department of Transportation, Washington, D.C., 2005.
5. Schweizer, T. Methods for counting pedestrians. In *The 6th International Conference on Walking in the 21st Century*. Walk21-VI "Everyday Walking Culture". Zurich, Switzerland, 2005.
6. Schwartz, W. and C. Porter. *Bicycle and Pedestrian Data: Sources, Needs, and Gaps*. Publication BTS00-02. Bureau of Transportation Statistics. U.S. Department of Transportation, Washington, D.C., 2000.
7. Dharmaraju, R., D.A. Noyce, and J.D. Lehman. An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicycles. In *The 71st ITE Annual Meeting*. Compendium of Technical Papers. Institute of Transportation Engineering, Chicago, IL, 2001.