Journal of Safety Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Motorcyclists' self-reported riding mileage versus actual riding mileage in the following year

Vicki Williams, [a,*] Shane McLaughlin, [a] Robert McCall, [a] Tim Buche [b]

[a] *Virginia Tech Transportation Institute, United States*
[b] *Motorcycle Safety Foundation, United States*

ARTICLE INFO

Article history:
Received 31 January 2017
Received in revised form 17 May 2017
Accepted 9 October 2017
Available online xxxx

1. Problem

Motorcycle crashes and fatalities are on the rise, and multiple transportation agencies have called for research into this issue. The motorcyclist fatality rate per 100,000 registered motorcycles in 2014 was six times that of the fatality rate for drivers of passenger cars (per 100,000 registered vehicles), and the fatality rate per motorcycle miles traveled was 27 times that of automobile miles traveled (National Highway Traffic Safety Administration, 2016). Many public and private organizations are interested in identifying causes of these crashes and related fatalities and injuries. One factor could be how capable riders are of estimating their own riding experience. The importance of "Self-Awareness" (including awareness of one's own skill level) is emphasized in the Motorcycle Safety Foundation (MSF) Basic Rider Course (2014). Further, in targeting training or considering survey responses, it is necessary to understand how indicative self-reported mileage is of actual future mileage. In addition, understanding the relationship between motorcyclists' estimated mileage and actual mileage is valuable because there continues to be some difficulty in obtaining accurate estimates of exposure (motorcyclist mileage) for measures such as crash and injury statistics as well as in efforts involving funding allocations, infrastructure planning, and financial forecasting (Lyon, Persaud, & Himes, 2017; Middleton et al., 2013).

Exposure estimates have been obtained via various methods, but fall short in terms of likely accuracy. Motorcycle registration does not provide a complete tally of riders due to unlicensed riders or licensing issued to individuals other than the actual rider. As indicated by the U.S. Department of Transportation (USDOT, 2015), the Motorcyclist Fatality Rate includes the "statistical issue" that the Federal Highway Administration (FHWA) likely underestimates the number of motorcycles on the road each year (supported by the finding from organizations such as the Motorcycle Industry Council that not all riders register their motorcycles). Annual motorcycle inspections rely on accurate odometer readings and faithful inspection scheduling. Observational and roadway detector equipment recordings of motorcyclist traffic flow present multiple difficulties related to collection protocol, location and timing choice, and sensor accuracy. Middleton et al. (2013) discuss some of these shortcomings, and offer various guidelines for calculating the accuracy of current methods used to report motorcycle traffic data. The authors provide detailed research methods and recommendations in terms of equipment and collection methods, and note that there are ongoing efforts to improve motorcycle traffic data. Their paper also presents the possibility of supplementing travel data through motorcyclist surveys, such as the National Household Traffic Survey (NHTS), origin and destination (O & D) surveys, and driver exposure surveys. If this additional source of motorcyclist mileage information (in the form of self-reported mileage) appears to be fairly accurate, it could be used to check or supplement mileage estimations collected through other methods.

As the first large-scale naturalistic instrumented motorcycle study to collect real-time mileage and self-reported mileage estimates, the MSF 100 study provides a unique data set from which to draw inferences about the characteristics of motorcyclist self-reported mileage and its application. Investigation of the data found in the study provides knowledge useful in answering multiple questions. Is collection of motorcyclists' self-reported mileage a useful method of checking or supplementing motorcycle travel data (which is in need of improved

* Corresponding author.
E-mail addresses: vwilliams@vtti.vt.edu (V. Williams), smclaughlin@vtti.vt.edu (S. McLaughlin), RMcCall@vtti.vt.edu (R. McCall), tbuche@msf-usa.org (T. Buche).

https://doi.org/10.1016/j.jsr.2017.10.004

accuracy)? Can we rely on rider self-reporting, especially reports of previous annual mileage, to estimate current or future mileage? If we survey riders about their previous mileage, can we make any predictions about the upcoming year? Would it be better to phrase this inquiry in terms of mileage during the most recent year, or would we most likely obtain a better estimate if we ask about the rider's overall annual average mileage (perhaps riders, for example, tend to perceive that they ride more now than in previous years)? Although motorcyclists are not expected to ride the same number of miles from year to year, any pattern in mileage estimates for previous years versus actual mileage on the road in the upcoming year can be informative in starting to uncover the actual relationship between estimated and actual rider mileage.

2. Method

The MSF 100 Motorcyclists Naturalistic Study was sponsored by the Motorcycle Safety Foundation and conducted by the Virginia Tech Transportation Institute (VTTI) to collect real-world riding data from riders in their natural day-to-day experiences while riding their own motorcycles. Individual rider participation in the study ranged from two months to two years. These 100 riders resided in California, Florida, Virginia, and Arizona, and both video data and motorcycle kinematic data were collected for every trip (defined as the time between key-on and key-off, during which the rider travels from one destination to another). These data were collected via unobtrusive instrumentation of a VTTI-developed data acquisition system (DAS) on each motorcycle, which continuously recorded five video views of the rider and the surrounding environment as well as motorcycle data such as GPS (Global Positioning System), acceleration, gyro, and brake activation. The full set of collected data incorporated over 366,000 miles of riding. Participants also completed various questionnaires prior to equipment installation, including riding exposure surveys, indicating riding habits to date such as annual mileage.

This paper explores motorcyclists' self-reported annual riding mileage and the actual amount of riding done within the study. For this evaluation, 91 riders (those who had been riding for at least one year before study enrollment and reported an annual mileage for that year) were considered. Self-reported annual mileage was recorded immediately preceding study participation, directly from the survey question "Approximately how many miles have you ridden a motorcycle on public roads in the past 12 months?" This sample of 71 males and 20 females were of various age groups and represented all three types of motorcycles, as indicated in Fig. 1.

Two approaches were used for determining the actual miles ridden. For 78 cases, starting and ending odometer readings were used. These

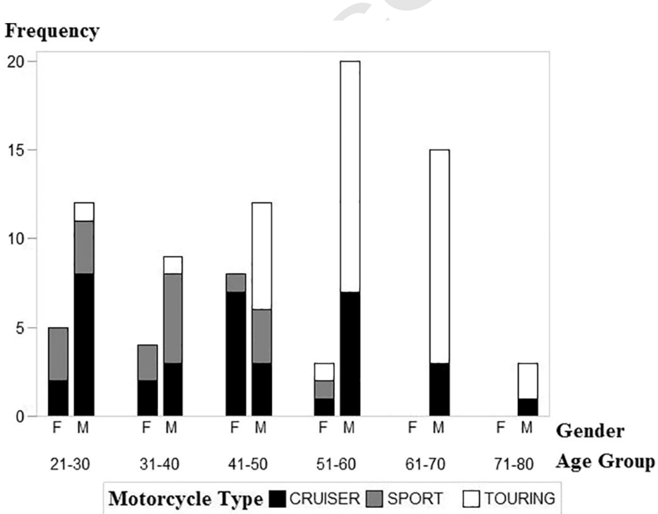

Fig. 1. Description of 91-Rider Sample.

readings were collected by technicians during equipment installation, and again during de-installation (at the close of participation). During the incorporation of the odometer readings for the data analysis phase, 13 cases were observed with recorded odometer reading that were missing or suspect (did not reasonably pair with the corresponding reading). In those 13 cases, integration of GPS-based speed and time data were used to calculate the distance traveled. Missing GPS values within a trip, including the period of time between DAS start and GPS signal acquisition, were replaced with the mean speed of that same trip. In the event that an entire trip was missing GPS speed, then the mean speed across all trips taken by that participant was used in conjunction with the trip duration to estimate mileage. However, the GPS dropout rate was not such that major adjustments were necessary, and thus total mileage estimates were not significantly affected. In general, the riders were not surrounded by high rise buildings or in weather that would interfere with GPS signal, and cellular signals were not used (the GPS was onboard with an external antenna).

All riding data (including odometer mileage and GPS-based calculations) were compared within riders to ensure the most accurate actual mileage was being used in the final analysis. The mileage for each rider was then translated to an adjusted annual riding mileage based on their study participation duration. At the beginning of the study, riders completed a survey which included a question about the number of months they tend to ride each year. Because the accuracy of self-reported data is not guaranteed, nor is it necessarily predictive of how many months overall they *would* have ridden during the study year, annualized mileage calculations were not based on any attempts at adjustment for riding season. However, each case of a reported abbreviated riding season was investigated to consider the potential effect of the riding season on annualized mileage calculations. The majority of riders who reported an abbreviated riding season participated for around a year and/or recorded very low mileage, so the riding season was either taken into account or was likely altered by a relatively small amount. Those riders falling outside of this category tended to ride nearly year-round, so any over- or under-estimation would also be relatively small.

During initial quality control efforts related to this MSF 100 study, one aspect of ensuring data integrity involved a video analysis sampling technique to prevent the inclusion of non-participant rider data files in the final data set. Although participants were informed that riding by anyone other than themselves was to be reported so applicable files could be deleted from the data set, this further analysis provided an extra safeguard against including non-participant trips. To verify that the rider was likely the consented study participant for all files included in MSF analyses (since viewing all of the more than 30,000 trip files was time- and cost-prohibitive), a video review sampling technique was conducted. This VTTI-developed tool, the Rapid Driver Identification (RDI) task, was used in the SHRP2 Naturalistic Driving Study, and is described fully for that application in McClafferty, Perez, and Hankey (2015). For the MSF review of each participant's selected files, two snapshots known to be the rider (one with and one without a helmet) were downloaded into the RDI system as a reference for comparison. Then snapshots from the file to be tested were gathered by dividing the trip into three segments of equal length (beginning, middle, and end) and using a face-detection algorithm to select a maximum of 12 snapshots from the trip (4 from each segment). At this point, through the video viewing tool, an analyst compared the known rider reference snapshots to the sample of actual file snapshots to determine whether the rider in the file was the participant. Appropriate quality control checks were also performed.

The goal of the rider identification task was to protect against a suspected non-rider contamination of 20% or more of each rider's trips (assuming that such contamination would significantly alter behavioral conclusions about that rider, such as substantially affecting collected mileage). Based on hypotheses testing and Bayes factors testing, if 6 or more out of a random sample of 20 of a rider's video files were contaminated (not the consented rider), then 20% or more

Please cite this article as: Williams, V., et al., Motorcyclists' self-reported riding mileage versus actual riding mileage in the following year, *Journal of Safety Research* (2017), https://doi.org/10.1016/j.jsr.2017.10.004