Journal of Safety Research 40 (2009) 197–201

Contents lists available at ScienceDirect

Journal of Safety Research

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

www.nsc.org

# Absence behavior as traffic crash predictor in bus drivers

Anders E. af Wåhlberg [a,*], Lisa Dorn [b]

[a] Department of Psychology, Uppsala University, P. O. Box 1225, 751 42 Uppsala, Sweden
[b] Department of Human Factors, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom

ARTICLE INFO

Available online 3 May 2009

*Keywords:*
Traffic crash
Accident
Bus driver
Absence
Exposure

ABSTRACT

*Problem:* Various indicators of health have been shown to be associated with traffic crash involvement. As general health is also related to absence from work, the latter variable may be more strongly related to crashes, especially for professional drivers. *Method:* Bus driver absence from work was analyzed in association with their crash records. Two British samples and one Swedish sample were used. *Results:* One of the British samples yielded fair correlations between crash record and absence, while for the other the effect was restricted to the first three months of driving. The Swedish data had effects in the expected direction but these were not significant. *Discussion:* The use of an indirect, overall measurement of health, may be a viable method for predicting the traffic crash involvement for professional drivers, although replications are needed in larger samples and other populations. *Impact on industry:* The use of absence records for the identification of at risk drivers would seem to be a simple and useful method for companies with major fleets, and it also shows the importance of promoting employee health and well being at work as a potential method of reducing the cost, not only of absenteeism, but also of crashes in company vehicles.



## 1. Introduction

Within traffic safety research, the attempts to predict individual differences in crash liability have received a fair amount of attention (see reviews by Arthur, Barrett, & Alexander, 1991; Clarke & Robertson, 2005; Lester, 1991). As most of the predictors used have been very specific, in that they probably are causally related to only a subset of crashes, effect sizes have generally been small. One set of predictors that have received little attention is health-related ones. Among those found to be associated with traffic crash record are obesity (Yamamoto, Akashiba, Kosaka, Ito, & Horie, 2000), sleep apnea (Barbe et al., 1998; Findley, Unverzagt, & Suratt, 1988; Young, Blustein, Finn, & Palta, 1997), and possibly exercise (Taylor & Dorn, 2006, 2008). For medical condition and treatment variables, correlations are generally very small, but reliable effects seem to have been found for diabetes (Cox et al., 2003), epilepsy (Lings, 2001), and various medications (Leveille et al., 1994; Hemmelgarn, Suissa, Huang, Boivin, & Pinard, 1997; and possibly Lagarde, Chastang, Lafont, Coeuret-Pellicer, & Chiron, 2005, although it was uncertain whether the effects were due to health problems as such or the medications to treat the ailments).

The interesting aspect of the previous research in the use of health-related variables as predictors of crash involvement is that they should all be associated with absence from work. As yet, research to show a link between absenteeism and traffic crash involvement for professional drivers (who have a higher exposure to crash risk) has not been demonstrated. Furthermore, if the wide ranging health conditions previously studied yield an effect on crashes as well as absence, then an overall measure of health, such as absence, should have a larger effect than these variables individually. It is therefore possible that absence from work may be a good predictor of traffic crash involvement.

Absence from work is in itself an unwanted behavior for companies; a link between absence and at-work crashes for professional drivers could be important information for managing at-work road risk.

Over 75 years ago, similar research on workplace injuries and absenteeism was reported for industrial workers (Farmer & Chambers, 1926, 1929), but the associations found can be suspected to be somewhat inflated, due to absences being caused by the injuries sustained (Verrier & Chevalier, 2007). However, in heavy vehicle (>10 tons) driver populations, such an effect is unlikely, due to the protection offered to the driver by the weight and sturdiness of such vehicles. This is especially true for bus drivers in city areas, where speeds are low and injuries to drivers are extremely rare (af Wåhlberg, 2002). Also, bus drivers have high levels of work-related health problems (Kompier & Di Martino, 1995; Long & Perry, 1985; Tse, Flin, & Mearns, 2006), making them a good population in which to study the proposed effect. The present study therefore sets out to investigate whether absence from work is associated with bus crashes for at work bus drivers. Of empirical interest is the notion of broad variables being better crash predictors than specific ones, but this is beyond the scope of the present study.

* Corresponding author. Tel.: +46 18 471 25 90; fax: +46 18 471 21 23.
*E-mail address:* anders.af_wahlberg@psyk.uu.se (A.E. af Wåhlberg).
*URL:* http://www.psyk.uu.se/hemsidor/busdriver (A.E. af Wåhlberg).

0022-4375/$ – see front matter 
doi:10.1016/j.jsr.2009.03.003

## 2. Method

### 2.1. Samples

Three sets of bus driver data were used for the present study; two from a major UK bus company and one from a small bus company in Sweden (Gamla Uppsalabuss (GUB)). Data on descriptive variables for these samples are shown in Tables 1 and 2.

In the UK, the operators in Sample 1 were newly hired and had taken part in a simulator-based study that required their crash and absence data to be collected for a period of 12 months after they began operational driving (Dorn, Stephens, & af Wåhlberg, submitted for publication). Therefore, these data had differing starting points and were not for the same time period for individual drivers. Traffic crash data were available for every three months of operational driving for each of the drivers.

Sample 2 consisted of bus drivers from two different locations within the UK (London and the North West of England, including North Wales and Manchester). Some operational and situational differences can be noted between these two groups. The London area has heavier traffic, more modern vehicles (and fewer different types), more CCTV surveillance on the vehicle, no fares collection by the drivers, and more supervision from managers. Data for these groups were added together across both samples due to small Ns. Both UK samples consisted of full time drivers only.

For the Swedish sample (sample 3), all drivers who had worked on full time schedules from 1999 to 2005 (as identified by their employee numbers) were included in the study, apart from those who worked less than 10 hours in any year (about two standard deviations from the mean for each year). Gamla Uppsalabuss (GUB) run all intra-city bus routes of Uppsala, the fourth largest town of Sweden, and has about 180 buses and 350 drivers at any time. A substantial number of the drivers work part time. Descriptive data for one year for Sample 3 are shown in Table 2.

**Table 2**
The descriptive data (means and standard deviations) for Sample 3 (Sweden).

| Number of subjects | Percent men | Age in years (mean/std) | Years of employment (mean/std) | Culpable crashes 1999-2005 (mean/std) | Hours worked 1999-2005 (mean/std) |
|---|---|---|---|---|---|
| 150 | 88% | 48.6/8.3 | 16.3/9.2 | 1.32/1.52 | 10 981/1902 |

Percent men, age, number of years with the company in 2003, and number of culpable crashes. Drivers with less than ten hours worked per year excluded.

### 2.2. Data

For samples 1 and 2, absence (number of days off work) and crash data were available from the employer for one year for sample 1 and 6 months for sample 2. All incidents resulting in damage and/or injury involving the bus are reported by the drivers (and therefore included in the presently used data), and culpability for these (including falls in the bus) are determined by the driver's manager and an insurance team, taking in such information as the driver's report, photographs, witness reports, and so forth. In the present study, two crash variables were used (All crashes and Culpable crashes) to test the hypothesis that absenteeism is associated with crashes. Theoretically, only crashes that the drivers have had some part in causing should be used as a criterion, as non-culpable incidents cannot be predicted (af Wåhlberg, 2008a,b; af Wåhlberg & Dorn, 2007, 2009).

However, it has been shown that the criterion used by this bus company to determine responsibility is somewhat too lenient (af Wåhlberg & Dorn, 2007). This means that the 'culpable only' variable will exclude a number of incidents that should rightly have been included, while the 'All' variable includes about 30% of incidents that are not the bus driver's fault and should therefore be excluded. As it was not possible to determine which crashes should have been excluded, both variables were included, as the "true" numerical value of culpable incidents is fully captured by neither variable.

For the Swedish drivers, number of hours spent at work and crash data were available for the period 1999-2005. Regarding culpability for crashes, the criterion used for categorization was shown to be approximately correct, classifying about 70% of crash involvements as at least partly the fault of the bus driver (af Wåhlberg & Dorn, 2007). Therefore, only one incident variable was used; culpable crashes. This consisted mainly of collisions (resulting in visible damage) with other buses and cars, but also injuries inside and outside the buses and hitting animals. Falls in the buses were not included unless they resulted in injury. Further descriptions of the bus company and their incident data can be found in af Wåhlberg (2002).

### 2.3. Statistical methods

The main aim of the study was to estimate the strength of the association between absence and traffic crashes. This can be expressed with many statistical methods, but Pearson correlation was chosen because it can be directly converted into percent explained variance, which is closely related to how well drivers can be correctly classified for crash involvement in a testing situation. The crash data used were not normally distributed, but this restricted variance only has the effect of limiting the size of the correlations (Glass & Hopkins, 1996). However, this effect is small: even when crash data with very restricted variance are used, the results follow a pattern that does not indicate that the correlations shrink faster than the mean of the samples used (af Wåhlberg & Dorn, 2009). Another indication of the robustness of the Pearson correlation against violations of its assumptions is that log transformation against skewness has no impact (Avolio, Kroeck, & Panek, 1985). Also, as shown by Gebers (1997), what statistical method is used has little impact upon the final result, in terms of explanatory power.

As a complimentary analysis, independent t-tests were used to bring out possible non-linear effects, where it was possible to dichotomize data in a meaningful way. ANOVA-type tests could not be used, as the crashes per hour variable was continuous.

No adjustments of alpha values for number of tests undertaken were implemented, as the interpretation of results are not based on any single values, but the overall picture.

### 2.4. Analyses

The analysis performed involved correlations being run between number of crashes for each driver for a specified time period and number of absence days for this person for the same time period. All

**Table 1**
The descriptive data (means and standard deviations) for samples 1 and 2;% men, age, years since car or PCV licensing, number of months with the company, number of crashes (all) and days of absence in the period given in the last column.

| Sample | Number of subjects | Percent men | Age in years (mean/std) | Years of holding a driving license (mean/std) | Months of employment (mean/std) | All crashes (mean/std) | Days of absence (mean/std) | Time period for data in this table |
|---|---|---|---|---|---|---|---|---|
| UK sample 1 | 176 | 86.7% | 37.6/10.0 | - | - | 0.59/0.83 | 1.03/2.43 | 3 months (after training) |
| UK sample 2 | 153 | 92.2% | 47.4/10.2 | 13.4/12.1 (PCV) | 100.3/105.7 | 2.99/2.63 | 1.84/0.89 | 12 months |