Transportation Research Part F 54 (2018) 349–356

Contents lists available at ScienceDirect

Transportation Research Part F

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

The effect of driver improvement interventions on crash involvement; has it been under-estimated?

A.E. af Wåhlberg

Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

ARTICLE INFO

Article history:
Received 19 October 2017
Received in revised form 14 February 2018
Accepted 22 February 2018
Available online 5 March 2018

Keywords:
Evaluation
Driver education
Driver training
Culpability
Intervention

ABSTRACT

The available evidence suggests that driver improvement interventions (with the aim to increase driver safety, most often by education or training) do not work. The average effect calculated in several meta-analyses is close to, and not always possible to distinguish from, zero, despite total samples sizes of several hundred thousand drivers. However, it is possible that all studies included in these meta-analyses have under-estimated the effect, due to a methodological error; all crashes have been used as dependent variable, instead of only those that the targeted drivers have caused. This error is expected to have considerably deflated the effect sizes, but it is not known how large this effect could be.

Using crash data for bus drivers in which culpability had been reliably established, a simple simulation was performed to determine the difference between using culpable and all crashes as an estimator of a safety effect. Using data for six years, calculations were made on single years. About ten percent of culpable crashes in each year were deleted to simulate a safety effect, where after the difference between the original and the simulated variable were calculated, using culpable only and all crashes in parallel. The effects using these two different kinds of datasets could then be compared and the under-estimation effect estimated.

Culpable crashes, as compared to all crashes, yielded larger differences in means between time periods, and smaller standard deviations. In between-subjects comparisons resulted in 15–30 percent larger effects for culpable crashes. Within-subjects calculations yielded larger but not as systematic effects.

The effect of driver improvement on crash involvement has been systematically under-estimated, as extremely few evaluation studies seem to have taken culpability for crashes into account. Therefore, new evaluations need to be undertaken, and/or old data re-analysed, to calculate a better estimate of the true effect of training and education in driving safety.

1. Introduction

One of the few problems that traffic safety researchers seem to agree upon is that driver improvement (which in this paper is taken to include all kinds of interventions where the aim is to reduce crash rates, usually in the form of training and education, for both problem and beginner drivers) does not have any proven safety effects (i.e. crash counts are not reduced), or that the effects are so miniscule as to be of no practical significance. Most of the many reviews and

E-mail address: aw@empirica.pm

https://doi.org/10.1016/j.trf.2018.02.027
1369-8478/

meta-analyses on this subject have reached this conclusion (Christie, 2001; Kaestner, 1968; Ker et al., 2005; Klein, 1966; Lonero, 2008; Lund & Williams, 1985; Masten & Peck, 2004; Mayhew & Simpson, 2002; Mayhew, Simpson, Williams, & Ferguson, 1998; McGuire & Kersch, 1969; Peck, 2011; Roberts, Kwan, & Cochrane Injuries Group Driver Education Reviewers, 2008; Strathman, Kimpel, & Leistner, 2007; Struckman-Johnson, Lund, Williams, & Osborne, 1989; Vernick et al., 1999; see Table 1 for some effect sizes), possibly with some caveat for possible effects of some specific sort of curriculum, and drink driving interventions (Wells-Parker, Bangert-Drowns, Mcmillen, & Williams, 1995). It can also be pointed out that most of the interventions included in these analyses were aimed at problem drivers (repeat violators and crashers) or young drivers (who have a higher risk than older drivers). The possibilities for achieving and finding an effect have therefore largely been maximised, at least in terms of statistics. Meanwhile, organisations that deal with transportation and safety have continued to educate and train drivers as if nothing had happened.

One of the problems in the evaluation of effects of driver improvement in traffic safety is the low variance of the accident variable. This yields very low power, so even if there is indeed an effect of training, this will be very difficult to detect. Recently, Peck (2011) calculated that we would need tens of thousands of drivers to achieve reasonable power (80% chance of detecting a ten percent reduction in crashes). Still, the number of studies on safety effects and the total number of subjects are rather large (see Table 1 for an overview of meta-analyses on this subject). However, several researchers have pointed out that the quality of most studies on improvement courses is doubtful (Lonero & Mayhew, 2010; Lund & Williams, 1985; Peck, 2011; Wundersitz & Hutchinson, 2006), in terms of not being randomized, having small sample sizes, using variables with doubtful validity etc.

It is well known that studies with doubtful methodology tend to report large effects (e.g. af Wåhlberg, Barraclough, & Freeman, 2016; Rongen, Robroek, van Lenthe, & Burdorf, 2013), and the evidence regarding driver education and training might therefore be even less positive than what might be believed, as whatever few positive effects have been found may be artefacts. The outlook for evidence in favour of driver training and education would therefore seem to be bleak.

The conclusion that driver improvement does not have a practically useful effect is thus based upon a fair amount of data, but the quality of this data can be questioned. The aim of the present paper is to present yet another problem inherent in evaluations of safety interventions, which might have systematically deflated the effect sizes in such studies.

First, it can be noted that a peculiar effect exists within driver improvement evaluation research; calculations on crashes and offences tend to yield different results (Struckman-Johnson et al., 1989), with the latter having larger effects. In the Masten and Peck meta-analysis (2004), the difference was 100 percent for the effect sizes (d .03 for crashes versus 0.06 for violations), which is the same as in Ker et al. (2005). Expressed in percent, the difference was smaller (28%) in the Masten and Peck study (−6.49% versus −8.28%), while Ker et al., did not report the results in percent. This effect could possibly be explained by differences in variance, although an actual differential effect on crashes and violations is possible. Offences and crashes correlate only about .18 (Barraclough, af Wåhlberg, Freeman, Watson, & Watson, 2016), and it is therefore very apparent that they to a large degree are due to different behaviours. Here, however, it will be suggested that part of the effect difference is due to a lack of validity of the crash variable.

Turning to how improvement evaluations are usually undertaken, there exists one methodological problem which has not been discussed in the literature, and which would impact negatively upon the effects found. This concerns the culpability for the crashes used in the evaluations. Usually, this is not taken into account, i.e. all crashes are used as dependent variable.

This is problematic, because driver training and education aim to change the accident-causing behaviour of the driver. It does not aim to reduce the exposure to risk from being hit by other drivers by no fault of their own. Therefore, if all crashes are used as the dependent variable, this variable will contain a fair amount of error variance, which reflects amount of exposure and not the quality of driving behaviour.

It has been proposed that in studies of individual differences in driving safety, only culpable crashes should be used as the dependent variable (af Wåhlberg, 2003, 2009), as this will yield a purer measure of the intended construct, and therefore larger effect sizes (as shown in af Wåhlberg, 2008, 2009). This logic can be extended to evaluations, as the basic problems

Table 1
Effect sizes for some driver interventions versus crashes in various meta-analyses. A positive d means a reduction in crashes.

Study	Type of population	Type of intervention	Dependent variable	N educated/control or k	Effect size (confidence intervals), type of statistic
Ker et al. (2005)	Average drivers	Remedial education	Crashes	180,563/54,380	0.98 (0.96–1.01) Risk ratio
Ker et al. (2005)	Average drivers	Advanced education	Crashes	35,539/12,222	0.99 (0.93–1.05) Risk ratio
Ker et al. (2005)	Average drivers	Remedial education	Injury crashes	34,186/23,263	1.17 (0.89–1.54) Risk ratio
Ker et al. (2005)	Average drivers	Advanced education	Injury crashes	1103/537	0.94 (0.74–1.20) Risk ratio
Masten and Peck (2004)	Problem drivers	Education, sanctions etc	Crashes	Total 1,640,000	0.030 (0.027–0.034) Cohen's d −6.49%
Roberts et al. (2008)	Pre-license drivers	Education	Crashes	Total 17,965	1.04 (1.02–1.05) Risk Ratio 1.01 (0.83–1.23) Risk Ratio 1.03 (0.98–1.09) Risk Ratio
Wells-Parker et al. (1995)	Drink/drive offenders	Education/counselling	Non-alcohol related Alcohol-related	k = 5 k = 10	−0.11 Cohen's d? 0.07 Cohen's d?