Accident Analysis and Prevention 75 (2015) 245–251

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Validation of the driver behaviour questionnaire using behavioural data from an instrumented vehicle and high-fidelity driving simulator

S. Helman*, N. Reed[1]

Transport Research Laboratory, Crowthorne House, Nine Mile Ride, Wokingham, Berkshire RG40 3GA, UK

ARTICLE INFO

Article history:
Received 14 January 2014
Received in revised form 10 December 2014
Accepted 10 December 2014
Available online 17 December 2014

Keywords:
Driver behaviour questionnaire
DBQ
Self-report scales
Instrumented vehicle
Driving simulator

ABSTRACT

Data from two previously published studies were used to examine the correlations between scores on the violation, error and lapse sub-scales of the driver behaviour questionnaire, and observed driving speed. One dataset utilised data from an instrumented vehicle, which recorded driver speed on bends on a rural road. The other utilised data from a driving simulator study. Generally in both datasets the DBQ violation subscale was associated with objectively-measured speed, while the error and lapse sub-scales were not. These findings are consistent with the idea that the DBQ is a valid measure of observed behaviour in real driving (its original intended use) and also in simulated driving. The fact that associations were the same in real and simulated driving lends further support to the relative validity of driving simulation. The need for larger and more focused studies examining the role of different motivations in different driving situations is discussed.

1. Introduction

Road safety research has made wide use of self-report scales to measure attitudes (e.g. Parker et al., 1996), behaviours (e.g. Reason et al., 1990) and safety outcomes such as accident involvement (Wells et al., 2008; Forsyth et al., 1995; Maycock et al., 1991). Such scales are cheap and easy to administer to large numbers of participants, and present a standardised way of gathering data on variables that are not always open to everyday observation. Much of the way that road safety researchers and policy-makers think about the links between attitudes, behaviours and safety outcomes is informed by studies that have made use of such scales.

Another method that has enjoyed widespread use in road safety research is driving simulation. Driving simulators have provided useful insights into the impacts on driving performance of alcohol (Mortimer, 1963), distraction (e.g. Burns et al., 2002; Reed and Robbins, 2008), social drugs (Sexton et al., 2000), fatigue (e.g. Philip et al., 2005) and new vehicle technologies (e.g. Hoedemaker and Brookhuis, 1998). Although not cheap to administer in the way that self-report measures are, simulators offer a number of other benefits for the investigation of road safety issues. Firstly, they enable potentially risky behaviours to be studied with no risk of harm to the driver or any third parties. Secondly, realistic, repeatable driving scenarios can be developed to elicit behaviours of interest and that could be unsafe, impractical or too costly to create in the real world. Thirdly, by design, simulators generate accurate, high frequency data about the driver (e.g. steering, pedals), the simulated vehicle (e.g. yaw rate, acceleration) and its position in the virtual world, which can be exploited by the experimenter.

No research method is free from limitations; neither self-report measures nor simulators are an exception. Given the extent to which such measures are used in road safety research it is useful to understand the scope of these limitations.

In particular, there is a fundamental criticism that has been levelled at both methods, related to the extent to which they can be taken as representative of 'real world' behaviour (see e.g. Evans, 2004, and af Wåhlberg, 2009). In this paper we present reanalyses of data from several existing studies that we hope will contribute to this debate. Specifically, we present data that directly assess the association between a widely used self-report scale of driver behaviour, behavioural data collected under quasi-naturalistic driving in an instrumented vehicle, and behavioural data collected during simulated driving. Before proceeding, it is useful to summarise some of the limitations and criticisms of self-report measures (and of driving simulation) in more detail.

1.1. Limitations of self-report measures

af Wåhlberg and his coworkers have been perhaps the most vocal in suggesting that self-report measures have serious

* Corresponding author. Tel.: +44 1344 770650.
E-mail addresses: shelman@trl.co.uk (S. Helman), nreed@trl.co.uk (N. Reed).
[1] Tel.: +44 1344 770046.

http://dx.doi.org/10.1016/j.aap.2014.12.008
0001-4575/

limitations (see af Wåhlberg, 2009, 2010). Some of these limitations are entrenched in the wider debate regarding the general scientific merit of subjective measures in fields concerned with human performance and behaviour (see Annett, 2002) and this debate certainly has relevance to road safety research (McKenna, 2002).

One specific criticism of interest here applies to the conditions under which self-report measures might be used as proxy measures for safety outcomes such as accidents; the literature is replete with examples of self-report scales that have claimed utility as proxies for the accident liability of individual drivers (see for example de Winter and Dodou, 2010). af Wåhlberg (2009, 2010),) has argued however that correlations between such measures and accidents may be due to common method variance (Podsakoff et al., 2003).

Common method variance refers to a variety of response biases that can occur when similar methods (e.g. self-report) are used to collect data on different outcome measures. One example of this is that people may wish to report in a congruent manner on their accidents and the behaviours or attitudes they exhibit, if they perceive that a study is investigating a link between the two (although see Parker and Manstead, 1996, for a counter-argument that other biases might lead to an *underestimate* of the link between these two variables). Another example of common method bias might be that respondents may exhibit simple tendencies to respond at specific points on self-report scales (for example near but not at the bottom of a scale) and if such scales are common across measures then spurious correlations may result.

It seems unlikely that common method variance is responsible for all findings linking driver individual differences to accident risk (for example see Wells et al., 2008 and Ivers et al., 2009 for similar predictive patterns from video-hazard perception test data and self-report scales onto self-report and police-recorded collisions). However it is noteworthy that in a recent meta-analysis (de Winter and Dodou, 2010) one of the most widely-used self-report measures (the driver behaviour questionnaire, Reason et al., 1990) was found to correlate only with self-reported (and not recorded) accidents (although the very low samples size from the studies found with recorded accidents might account for this). On the basis of such findings, it can be argued that if road safety researchers wish to benefit from the relatively cheap and efficient data collection that self-report measures permit (in their search for individual difference variables that are important for road safety) then the extent to which such measures can be shown to predict accidents outside of conditions in which common method biases might exist needs to be established.

One potential counter-argument to this position is that many scales are not constructed as explicit predictors of accidents; they are often constructed as self-report measures of *behaviours* that are in turn known or strongly suspected to be good predictors of accident propensity (for example in the case of speed see Aarts and van Schagen, 2006). Logically however the same arguments regarding common method biases apply in this case; if self-report measures are correlated with other self-report measures (rather than objective measures) then common-method variance might again be to blame for correlations observed. af Wåhlberg (2010) notes that very few self-report measures of driver behaviour have been subjected to any validation using objective measures of the behaviours in question.

One way around this limitation is to use behaviour observed in naturalistic driving as the outcome measure against which to validate the self-report measure, something that is becoming possible with the increased functionality observed in in-vehicle data recorders. Simons-Morton et al. (2013) report a recent study in which this was done. Two self-report measures were validated against accelerometer data from instrumented vehicles driven by 42 newly licensed teenage drivers. The Checkpoints risky driving scale (C-RDS) and the DULA dangerous driving index (DDDI) were the self-report measures used. Simons-Morton et al., 2013 found significant and moderate correlations between the C-RDS and the overall measure of objective risk (based on accelerometer data from the first 18 months of driving). For the DDDI the picture was slightly less clear-cut, with only the 12 items constituting the 'risky driving' subscale being significantly and consistently correlated with the objective measure of risk.

One objective of this paper is to examine the external validity of a self-report measure of behaviour by making use of some objective driving behaviour data (in this case from a previously-completed instrumented vehicle study). The self-report measure used is the driver behaviour questionnaire (Reason et al., 1990). The DBQ is probably the most widely used self-report measure in the road safety literature and there has been considerable debate about its validity as a predictor of accidents. This debate is summarised in the meta-analysis published by de Winter and Dodou (2010) and subsequent commentaries and counter-commentaries on this article in the Journal of Safety Science from the original authors, and af Wahlberg and coworkers (af Wåhlberg and Dorn, 2012; de Winter and Dodou, 2012a,b,c; af Wåhlberg et al., 2012a,b). We believe that this debate is enriched by consideration of whether the DBQ has validity as a measure of real-world driving behaviour and therefore in this study, we examine the correlation between the DBQ subscales (violations, errors, and lapses) and speed choice, seeking to add to the limited existing literature on this topic.

Existing literature, to which we hope to add, generally supports the assertion that the DBQ (in particular the violations subscale, which comprises items that suggest an aggressive, fast driving style) possesses some validity as a proxy for various risky driving behaviours. Early work using the DBQ for example showed that on-road observed speeds correlate with scores on the DBQ violations subscale, even when age and mileage were taken into account (Quimby et al., 1999a,b,b; see also de Angeli et al., 1996). Some recent evidence also exists on this. Åberg and Wallén-Warner (2008) for example observed a direct correlation between the DBQ violations subscale and logged speeding in a trial of an intelligent speed adaptation (ISA) technology. Zhao et al. (2012) also showed that naturalistic driving speeds were associated with DBQ violations scores, but not with the other DBQ subscales (errors and lapses). Metz et al., (2013) demonstrated a positive correlation between DBQ violations and two speed measures in an instrumented vehicle study (mean speed and time greater than 20 km/h above the speed limit. Note however that Underwood (2013) did not find any significant correlation between the DBQ violations subscale and recorded on-road speed in an instrumented vehicle in novice drivers.

Links have also been observed between the DBQ violations subscale and observed violation outcomes, both prospectively and retrospectively. Palamara and Stevenson (2003) used a prospective version of the scale in a sample of 1277 Perth drivers obtaining their first licence, and showed that intention to commit violations was associated with speeding infringements in the first year of driving. González Iglesias and Gómez Fraguela (2010) found that those drivers who had committed driving offences (recorded by the police) scored more highly on the DBQ violation subscale.

We also take the opportunity to comment on another area of road safety research in which there is debate about validity–the use of driving simulators. As well as making use of instrumented vehicle data therefore we also make use of data from a driving simulator, in order that we may examine the nature of links between these data and the DBQ. Again there is some existing