INVESTIGATING THE CALIBRATION ABILITY OF YOUNG NOVICE DRIVERS RELATIVE TO EXPERIENCED DRIVERS: PRELIMINARY FINDINGS FROM A DRIVING SIMULATOR STUDY

Eve Mitsopoulos, Thomas Triggs & Michael Regan
Monash University Accident Research Centre
Building 70, Monash University, Victoria 3800, Australia
Ph.: +61 3 9905 1900; Fax: +61 3 9905 4363
Email: eve.mitsopoulos@general.monash.edu.au

ABSTRACT

The over-representation of young novice drivers in road crash statistics is a major public health concern. It is believed that errors in calibration are a contributory factor in young novice drivers’ crash involvement. Calibration in driving can be defined as the ability to match task demands to one’s own driving capabilities. To date, much of the research suggesting that young drivers are miscalibrated relative to experienced drivers is indirect or relies exclusively on the outcomes of subjective measures. Direct behavioural evidence is lacking. The current paper describes the method of, and lists some preliminary findings from, a driving simulator study which aims to examine objectively whether young novice drivers are indeed miscalibrated relative to experienced drivers in particular driving scenarios. Of the two driving situations under study, car following and gap acceptance at right hand turns, preliminary analysis of a subset of the data, revealed a difference between experienced and novice drivers in the case of car following only. The outcomes of this research will serve to better inform the design of countermeasures which aim to minimise the incidence of young driver crashes due to deficiencies in calibration ability.

KEYWORDS

Young novice drivers, calibration, driving skill, driving experience, driving simulator, road safety, car following, gap acceptance

INTRODUCTION

It is well recognised in the road safety community that young novice drivers are over-represented in motor vehicle crashes, and that this is true for all motorised countries. Moreover, from a public health perspective, motor vehicle crashes constitute one of the leading causes of death among young people. Much research in the area of young driver crash involvement has been instrumental in identifying some of the reasons behind young novice drivers’ high crash propensity. A useful classification distinguishes between three overarching reasons: skill/inexperience, age/youthfulness, and exposure (Drummond, Triggs, Macdonald & Bowland, 1993).

In the skill area, it has been proposed that errors in calibration contribute to young novice driver’s crash involvement (e.g., Triggs, 1994). Calibration in driving can be defined as the ability to match the demands of the driving task to one’s own capabilities as a driver. Calibration can be thought of as a meta-cognitive skill in that it involves an accurate awareness of the demands, both static and dynamic, imposed by the traffic system and accurate self-insight into one’s capabilities to meet those demands. Moreover, calibration requires comparison between capabilities and demands, to determine whether there is an undesirable mismatch between them which necessitates appropriate modification to one’s driving behaviour. That is, being well-calibrated involves being able to accurately judge whether, and how, driving behaviour needs to be modified by the individual driver so that he/she can effectively manage the demands imposed by the situation and, therefore, minimise the opportunity of a crash.
To date, much of the research suggesting that young drivers are miscalibrated relative to experienced drivers is indirect or relies exclusively on the outcomes of subjective measures (e.g., Finn & Bragg, 1986; Matthews & Moran, 1986). Direct behavioural evidence is lacking. Part of the problem may lie in finding or developing an appropriate experimental technique for adequately identifying and objectively measuring calibration ability in driving. The current paper describes the method of, and lists some preliminary observations from, a driving simulator study which aims to examine objectively whether young novice drivers are indeed miscalibrated relative to experienced drivers in each of two driving situations: car following and gap acceptance at right hand turns. Essentially, the study involves relating self-assessed performance to actual performance for the two driving situations separately that each vary in their level of demand. Differences in calibration ability between novice and experienced drivers can then be identified. In this approach, self-assessed performance is derived from responses to a series of simulator drives experienced in “autopilot” mode. In autopilot mode, participants do not have any control over the simulator vehicle. Participants make judgements about how likely they would be to perform a given behaviour if they were in control of the simulator vehicle. Actual performance is derived from a set of drives, which are similar to those experienced in autopilot mode, however, participants control the simulator vehicle themselves. A measure of performance, taken from when the participant was in control of the simulator vehicle, is then related to the corresponding judgement made when the participant was not in control of the vehicle. Comparisons are then made between novice and experienced driver groups for evidence of any discrepancies between the groups. This experimental technique draws on that used to measure calibration in domains outside of driving and road safety (e.g., Lichtenstein, Fischhoff, & Phillips, 1982; Keren, 1991).

METHOD

Experimental Design

For each of the car following and gap acceptance tasks, the experimental design was a mixed design with two independent variables. For both tasks, the between subjects factor was driver experience group: novice and experienced. For the car following task, the within subjects factor was following distance (metres). Five different levels of following distance were employed: 12, 23, 34, 45 and 56 metres. At a speed of 60 km/h, these represented time headways ranging from 0.7 to 3.4 seconds. For the gap acceptance task, the within subjects factor was gap size (seconds). Five different levels of gap size were used: 3.0, 4.1, 5.2, 6.3 and 7.4 seconds. The choice of levels within each of the car following and gap acceptance tasks was determined on the basis of the outcomes of previous research (e.g., Wennell & Cooper, 1981; Fairclough, May & Carter, 1997; Taieb-Maimon & Shinar, 2001; Cooper & Zheng, 2002) and, in particular, on the findings of a series of pilot tests. The aim here was to select a range of levels that, at the extremes, would be relatively easier to judge as turn/not turn or follow/not follow at a shorter distance, than levels in between, thus reflecting differing levels of task demand.

Participants

A total of 63 drivers took part in the study. Thirty-one participants (17 males and 14 females) belonged to the experienced driver group and the remaining 32 participants (18 males and 14 females) comprised the novice driver group. An additional five participants (7.4% of the total sample), two novice and three experienced, withdrew from the study due to simulator discomfort.

Participants in the experienced group ranged in age from 30.4 to 51.3 years (mean = 41.8 years; SD = 5.9), while participants in the novice group were aged between 18.1 and 21.7 years (mean = 19.2 years; SD = 1.1). All participants in the novice driver group held a

---

1 The minimum licensing age in Victoria, Australia is 18 years.
current Victorian Probationary\(^2\) car driver’s licence and had held that licence for no more than six months (mean = 3.0 months; SD = 1.6). Moreover, other than a Learner’s Permit, these drivers had not held another licence prior to obtaining their current Probationary licence. Participants in the experienced driver group were all holders of a full car driver’s licence and had accrued between 12.4 and 33.3 years of driving experience (mean = 23.1 years; SD = 5.9). All participants reported having normal or corrected-to-normal vision. Individuals with epilepsy or who were prone to motion sickness were not eligible to take part in the study.

Participants were primarily staff and students from Monash University. They were recruited through several means, including advertisements posted on University electronic notice bulletins and through posters and flyers placed or distributed at several locations around the Clayton campus of the University. All participants provided written informed consent prior to taking part in the study.

**Apparatus**

The Monash University Accident Research Centre’s advanced driving simulator was the main experimental tool used in the study. It comprises a Holden Calais sedan with normal interior fit out and an automatic transmission. The vehicle is surrounded by two projection screens: one screen is located to the front of the vehicle and one to the rear. From the driver’s eye the front screen provides a field of view subtending angles of approximately 180 degrees horizontally and 40 degrees vertically. The rear screen provides a field of view subtending angles of approximately 60 degrees horizontally and 40 degrees vertically. Four projectors, three front facing and one rear facing, convey images of pre-programmed traffic scenes onto the screens. (However, the rear facing projector and, therefore, rear screen, was not used during the current study.) The displayed images are generated by a Silicon Graphics Onyx computer, which updates the images at a rate of 30 Hz. In addition, the simulator vehicle is mounted on a motion platform, which is intended to provide the driver with some sensation of road feel. A quadraphonic sound system is utilised to provide the driver with realistic traffic sounds, which include engine noise and horn blasts.

The experimenter runs and monitors driving simulations from the control room, which is a separate room located adjacent to the simulator room. A two-way communication system between the control room and the simulator vehicle enables the experimenter and participant to communicate as required. Video monitors located in the control room and connected to cameras mounted inside the vehicle allow the experimenter to monitor participants for signs of discomfort and also to monitor the status of the simulations while they are in progress.

**Tasks and Procedure**

Participation in the study involved two sessions conducted on an individual basis and held one to two weeks apart. Session 1, the “Autopilot session”, lasted for about one-and-a-half hours, while session 2, the “Manual session”, took approximately one hour to complete.

**Autopilot session**

The Autopilot session was structured as follows. First, participants completed questionnaires designed to gather information on their basic demographic profile and driving experience, and to assess participants’ state of well-being prior to driving the simulator. In turn, participants completed the Familiarisation drive. The primary function of this drive, which lasted for approximately three minutes, was to acquaint participants with the various elements of the virtual driving environment and with the control dynamics of the simulator vehicle. The drive occurred on a two-lane residential road with a speed limit of 60km/h and some occasional oncoming traffic. Participants were asked to drive normally and to practice some of the basic manoeuvres that were pre-requisites for later drives (e.g. accelerating, decelerating, braking, and maintaining 60km/h). Towards the end of the drive, participants were asked to make a

\(^2\) Newly licensed drivers in Victoria, Australia hold a Probationary licence for three years before graduating to a full licence.
right-hand turn. No vehicles were present in the oncoming lane as this manoeuvre was performed. The Familiarisation Drive was performed twice in succession: on the first occasion, the experimenter sat in the passenger seat next to the participant, while on the second occasion, the experimenter was seated in the control room. In each case the experimenter provided participants with feedback and assistance as required.

Following the Familiarisation drives, participants were asked to undertake the Autopilot trials. These comprised 24 drives, each one approximately 45 seconds in length, experienced in “autopilot mode”. All drives occurred on a two-lane residential road with a speed limit of 60km/h and some occasional oncoming traffic. Participants were given the following instructions:

You will experience each of these drives in “autopilot mode”. What this means is that you will be in the driver’s seat of the simulator vehicle, but you will not have any control of the vehicle. The car has been programmed to drive itself… Each drive will pause at a certain point. This pause marks the end of the drive. At this point your task will be to make a decision about your likelihood of performing a particular behaviour if you were in control of the simulator vehicle…

Half of the drives depicted a car following episode, while a right hand turn at an unsignalised intersection (gap acceptance) characterised the remaining drives. During the car following drives, participants observed a car travelling at a certain distance in front of the simulator vehicle; after a few seconds the drive paused and participants were required to make a judgement about how likely they would be to follow the lead car at a shorter distance if they were in control of the simulator vehicle. In each case, participants made their judgement by placing a line at a point along a scale. The scale ranged from 0% (absolutely no chance that I would follow at a shorter distance) to 100% (absolutely certain that I would follow at a shorter distance). During the right hand turn drives, participants observed the simulator vehicle approach an intersection; at the same time four additional vehicles approached the same intersection from the opposite direction in the adjacent lane. The drive paused and participants’ task was to judge how likely they would be to turn between the black oncoming vehicle and the white oncoming vehicle if they were in control of the simulator vehicle. The black and white vehicles were the second and third vehicles, respectively, in the series of oncoming vehicles. In each case, participants made their judgement by placing a line at a point along a scale. The scale ranged from 0% (absolutely no chance that I would turn) to 100% (absolutely certain that I would turn).

Participants made each of their judgements in a booklet, which they kept with them in the simulator vehicle. Each page was restricted to a single judgement; participants were encouraged to make each judgement immediately after the given drive had paused and, once they had made their judgement, to turn the page in the booklet and to prepare for the next drive. Participants completed four Autopilot drives for practice. The first and second drives depicted a right hand turn event with a target gap of 3.5 and 6.9 seconds, respectively. The third and fourth drives were characterised by a car following event with a target following distance of 18 and 50 metres, respectively. Presentation of the experimental Autopilot drives followed. These comprised 10 drives: five car following (12, 23, 34, 45 and 56 metres) and five right hand turn (3.0, 4.1, 5.2, 6.3 and 7.4 seconds). Each drive was presented twice, giving 20 drives in total. The 20 experimental drives were presented to each participant in one of four pseudo-random presentation orders. Each order was administered a similar number of times within each of the experienced and novice driver groups, and consequently, between the two groups. The Autopilot session concluded with participants completing the questions on current well-being a second time, to gauge any changes in participants’ well-being following their time in the simulator.

**Manual session**

The Manual session followed a simulator structure to that of the Autopilot session. First, participants answered the questions on current well-being. Second, participants completed the Familiarisation drive twice – following the exact procedure used in the Autopilot session. Third, participants were asked to undertake the Manual drives. These comprised 12 drives,
which each lasted for approximately two minutes. The Manual drives differed from the Autopilot drives in two ways. In the Manual drives, participants had complete control of the simulator vehicle, that is, participants were asked to drive the simulator vehicles themselves. In addition, each Manual drive comprised both a car following event and a right hand turn at an unsignalised intersection.

All Manual drives occurred on a two-lane residential road with a speed limit of 60km/h and some occasional oncoming traffic. Participants were instructed to drive as they would in the real world taking into consideration the speed limit and other vehicles. Participants were also instructed not to overtake any vehicles and that the only time they were permitted to move out of their lane was when they were performing the right hand turn.

The right hand turn event was similar in appearance to that in the Autopilot drives in that, as the participant in the simulator vehicle approached the critical intersection, a series of four vehicles approached from the opposite direction in the adjacent lane. Participants were instructed to turn right at the first opportunity where they felt that they had enough room in the stream of oncoming traffic to turn safely. The positioning of the oncoming vehicles relative to the simulator vehicle at the intersection was such that participants effectively had only two options: to turn between the second (black) and third (white) vehicles or to wait until all four vehicles had passed the intersection before turning.

The car following event occurred in two phases. The first phase commenced when a vehicle positioned on the left hand side of the road merged into the simulator vehicle’s lane, ahead of the simulator vehicle. For several seconds, the lead vehicle travelled ahead of the simulator vehicle maintaining a pre-determined distance. Participants were instructed to maintain a speed of 60 km/h during this time. A single horn sound signalled the commencement of the second phase of the car following event. Once the horn sounded, the participants’ task was to travel at whatever speed was necessary in order to get to the following distance that they wished to adopt – a following distance that participants’ felt was safe for them. Participants were required to maintain their target following distance for several seconds after which they would hear a horn sound twice. At this time, the lead vehicle accelerated and drove into the distance signalling the end of the car following event. Each drive concluded shortly after the participant had completed both events.

Participants first completed two Manual drives for practice. The gaps/following distances used in the Autopilot practice drives were also used in the Manual practice drives. The first practice drive was composed of a right hand turn event (3.5s) followed by a car following event (50m), while the second practice drive was made up of a car following event (18m) followed by a right hand turn event (6.9s). Presentation of the experimental drives followed. The experimental drives were made up of two sets of five drives. In one set the right hand turn event occurred first followed by the car following event, and in the other set the car following event occurred first followed by the right hand turn event. Each gap (3.0, 4.1, 5.2, 6.3 and 7.4s) and following distance (12, 23, 34, 45 and 56m) level was represented once in each set of experimental drives. Accordingly, across the two sets, each participant was exposed to each gap and following distance twice. To produce the drives within each set, gaps and following distances were paired pseudo-randomly. The ten experimental drives were presented to each participant in one of four pseudo-random presentation orders. Each order was administered a similar number of times within each of the experienced and novice driver groups, and therefore, across the two groups.

The Manual session concluded with participants completing the questions on current well-being a second time, to gauge any changes in participants’ well-being following their time in the simulator. Participants also completed a further set of questions pertaining to their driving experience.
RESULTS

The primary analysis deriving from this study constitutes the generation of calibration curves for each of the gap acceptance and car following tasks for comparison between the driving experience groups. Essentially, this involves matching the judgement information obtained in the Autopilot sessions with the corresponding performance data collected as part of the Manual sessions. At the time of writing, however, data pertaining only to the Autopilot sessions were available. Accordingly, the remainder of this paper provides a preliminary analysis of these data for both the car following and gap acceptance tasks.

A similar set of analyses was undertaken for each task. As a starting point, participants’ judgements were categorised into several groups. This was done separately for each driving experience group and for each following distance, in the case of the car following task, and each gap size, in the case of the gap acceptance task. Responses were categorised into the following five groups, representing differing levels of likelihood of performing the behaviour in question: Low (0 to 19% likely); Low-Medium (20 to 39% likely); Medium (40 to 59% likely); Medium-High (60 to 79% likely); and High (80 to 100% likely). These data are presented in Figures 1 and 2 for the car following task and Figures 3 and 4 for the gap acceptance task. To maximise power, data were combined across the two trials in which participants were exposed to the same gap or following distance. This was justified for the current analyses, since, in general, no significant differences in median likelihood were found between the two trials of a given drive type.

For each task, a series of chi square analyses was undertaken, where, in every case, judgement response category was the dependent variable. First, novice and experienced drivers’ responses were treated separately in order to examine differences in pattern of responding between different levels of the following distance variable and also the gap size variable. While the responses for each level of the task were provided by the same group of participants, these responses nevertheless can be considered as independent observations, thus justifying the use of chi square for this first set of preliminary analyses. Second, for a given following distance or gap size, novice and experienced drivers’ responses were compared to ascertain any differences in response pattern depending on driving experience level. In conducting the analyses and reporting their outcomes, the following considerations were made as recommended by Siegel and Castellan (1989). First, where the degrees of freedom exceeded one and more than 20% of the cells had an expected frequency of less than five, adjacent categories were collapsed. Second, in cases where the degrees of freedom was equal to one and no cells had an expected frequency of less than five, chi square corrected for continuity was reported, otherwise the result of Fisher’s exact test was given.

Car Following

Figures 1 and 2 show that, for both the novice and experienced groups, respectively, the pattern of responses varied depending on the following distance. This difference was significant for the novice group ($\chi^2 (16) = 270.76, p<0.01$) and also for the experienced group ($\chi^2 (16) = 186.63, p<0.01$). As expected, for the smallest following distance, 12 metres, the majority of participants responded that their likelihood of following the lead car at a shorter distance was Low. The proportion of Low responses decreased as a function of following distance. The largest following distance, 56 metres, showed the reverse pattern, with the majority of participants responding that their likelihood of following the lead car at a shorter distance was High. Not surprisingly, the proportion of High responses increased as a function of following distance. Greater balance in the proportion of cases across response categories seems most apparent for the mid-range following distance of 34 metres in the case of the novice drivers and 45 metres in the case of the experienced drivers.

A series of analyses was conducted within adjacent pairs of following distances to determine in general terms whether there were any changes in response patterns across increments in following distance. Indeed, a significant change was found between each pair of judgement categories for both the novice and experienced groups.
Comparisons between the novice and experienced groups for each following distance separately, revealed significant differences in the pattern of responding between the two groups at all following distances with the exception of the smallest of 12 metres ($p = 0.12$, Fisher’s exact test). When faced with a following distance of 12 metres, participants felt that their likelihood of following the lead vehicle at a shorter distance was effectively low; a feeling that was independent of driving experience. At 23 metres, however, the pattern of responding varied significantly between the two groups ($\chi^2 (2) = 8.60, p<0.05$). In particular, while the likelihood of following at a shorter distance was reported to be Low for 79% of the cases in the experienced group, this figure was considerably lower for the novice group at 56%. Conversely, while the likelihood of following at a shorter distance was Medium, Medium-High or High combined for only 7% of cases in the experienced group, the corresponding proportion was observably higher for the novice group at 22%. A similar story could be said in the case of 34 metres for which there was also a significant difference ($\chi^2 (4) = 19.30, p<0.01$). Relative to the experienced drivers, the proportion of cases for the novice drivers where the likelihood of shorter following was Low was less than half. Moreover, while the majority of experienced drivers were still responding Low, the majority of novice drivers were responding Medium to Medium-High (55%).
The pattern of responding between the novice and experienced groups was also significantly different at 45 metres ($\chi^2(4) = 9.47, p=0.05$). While the majority of cases for the novice group occurred in the Medium-High to High categories (63%), this was not the case for the experienced group; rather, for the experienced group, a similar proportion of cases were reported in the Low to Low-Medium categories (34%) as in the Medium-High to High categories (39%). A significant difference in response pattern between groups was also yielded at 56 metres ($\chi^2(3) = 19.91, p<0.01$). In particular, the likelihood of following at a shorter distance was reported to be High in observably fewer cases for the experienced group than for the novice group. Conversely, the proportion of cases that fell into lower likelihood categories was greater for the experienced group than for the novice group. In summary, these results indicate that, relative to the experienced group, the novice group responded with higher certainty that they would follow at a shorter distance. This was true for all of the following distances under study, with the exception of the shortest.

**Gap Acceptance**

As can be seen in each of Figures 3 and 4, participants’ responses to the right hand turn judgement questions varied depending on the gap size presented in a given drive. Overall, this variation was found to be significant for the novice group ($\chi^2(16) = 200.79, p<0.01$) and also for the experienced group ($\chi^2(16) = 216.74, p<0.01$). The pattern of responding yielded for gap acceptance largely appears to mirror that observed for car following. Specifically, presentation of the smallest gap size of 3.0 seconds gave a Low likelihood of turning in the majority of cases, while the largest gap size of 7.4 seconds yielded a High likelihood of turning rating in most of the cases at that level. Further, while cases in the Low category decreased with each increment in gap size, cases in the High category, in general, increased. Finally, greater balance across the response categories can be observed for the gap sizes in the middle of the range, namely at 5.2 and 6.3 seconds. Comparisons within adjacent pairs of gap sizes confirmed that, in general, there was a significant change in response pattern with increments in gap size. There was one exception, however; for the novice group only, there was no significant change in the response pattern from 5.2 to 6.3 seconds ($\chi^2(4) = 5.51, p=0.24$).

![Figure 3](image-url)

**Figure 3.** Novice participants’ responses to the Autopilot right hand turn drives as a function of gap size and judgement response category
Comparisons between the novice and experienced groups for each gap size separately, revealed a very different pattern of outcomes to those which came out of the corresponding car following set of analyses. Essentially, there was no significant difference in pattern of responding between the novice and experienced groups at each of 3.0 seconds (p=0.44, Fisher’s exact test), 5.2 seconds ($\chi^2 (4) = 5.23, p=0.27$), 6.3 seconds ($\chi^2 (3) = 0.90, p=0.83$), and 7.4 seconds ($\chi^2 (4) = 7.57, p=0.11$). At 4.1 seconds, however, a trend towards significance was observed ($\chi^2 (2) = 5.11, p=0.08$). This was most likely to have been driven by a smaller number of cases in the Low response category for the novice drivers relative to the experienced drivers, and perhaps to some extent, by a larger number of cases in the Medium, Medium-High and High categories combined for the novice drivers (13%) compared with the experienced drivers (3%). However, taken overall, these outcomes lend very little, if any, support for a difference in reported likelihood of turning between the novice and experienced groups.

**DISCUSSION**

The purpose of this paper was to describe the method of, and to report on some preliminary observations from, a driving simulator study which aims to examine objectively whether young novice drivers are miscalibrated relative to experienced drivers. The study focussed on two driving situations: car following and gap acceptance at right hand turns. For each situation, the primary analysis will involve relating self-assessed performance, obtained in the Autopilot sessions, to actual performance, obtained in the Manual sessions. Comparisons can then be made between novice and experienced driver groups for evidence of any discrepancies between groups.

The current paper presented preliminary findings from the Autopilot sessions only, as data from the Manual sessions were not yet available at the time of writing. While analysis of the Autopilot data alone cannot provide evidence of miscalibration of one driver group relative to the other, these data, nevertheless, provide some interesting information pertaining to participants’ pattern of responding in terms of their perceived likelihood of performing a particular behaviour as a function of task demand. Analysis of the autopilot data revealed that, as might be expected, response patterns concerning likelihood of performing a particular behaviour varied depending on the level of task demand. This was the case for both the gap acceptance and car following tasks and also, for both the novice and experienced driver groups. However, differences between driving experience groups were only observed for the car following task with the novice group responding with higher certainty that they would follow at a shorter distance than the experienced group; a finding which was true for all of the
following distances under study, with the exception of the shortest. It is envisaged that the outcomes of this research will serve to better inform the design of countermeasures which aim to minimise the incidence of young driver crashes due to deficiencies in calibration ability.

ACKNOWLEDGEMENTS

The authors wish to thank the following individuals for their technical contributions during the course of the study: Ashley Verdoorn, Jason Rubens, and Nebojsa Tomasevic. We also acknowledge the financial support provided through the Monash University Accident Research Centre’s Doctoral Student Research Fund.

REFERENCES


