A SIMULATOR COMPARISON OF RIDING PERFORMANCE BETWEEN NEW, RETURNED AND CONTINUING MOTORCYCLE RIDERS

Mark Symmons & Christine Mulvihill
Monash University
Clayton, Victoria, Australia
Email: mark.symmons@monash.edu

Summary: Motorcycle riders constitute an increasing proportion of deaths and injuries due to road crashes. Further, older riders represent an increasing proportion of motorcycle riders, with particular concern for those who have returned to riding after a substantial break (returned riders). These riders seem reluctant to undertake any sort of refresher training. In a motorcycle simulator task continuing riders performed best, followed by newly licensed riders, and then returned riders, though that result did not extend to crashes that occurred in the simulator.

INTRODUCTION

In Australia, motorcycle riders constitute an increasing proportion of deaths and injuries due to road crashes (BITRE, 2010). This trend likely reflects an increase in the number of older motorcyclists, including those taking up riding for the first time (new riders) or returning to riding after a substantial break (returned riders). While the over-representation of newly licensed motorcyclists in crashes is relatively well documented, less is known about the nature and extent of crashes involving returned motorcyclists.

Survey research has identified a number of factors that might place returned riders at a higher crash risk compared to new and continuing riders. Returned riders are less likely to undertake training (Haworth, Mulvihill & Symmons, 2002; Haworth & Mulvihill, 2005; Mulvihill, Symmons & Liu, 2010; Sexton, Baughan, Elliot & Maycock, 2004), more likely to ride a motorcycle that is of a higher engine capacity than they did in the past (Haworth, Mulvihill & Symmons, 2002); more likely to be riding as part of a group at the time of a crash (Mulvihill & Symmons & Liu, 2010); and more likely to ride in rural areas and for recreation (Haworth, Mulvihill & Symmons, 2002; Haworth & Mulvihill, 2005; Mulvihill, Symmons & Liu, 2010; Sexton, Baughan, Elliot & Maycock, 2004). However, these findings are not sufficient on their own for developing a program to reduce returned rider crash involvement.

Motorcycle simulators have been used to train and assess novice riders (see Vidotto, Bastianelli, Spoto, Torre & Sergeys, 2008). They have also been employed for research purposes. For example, Liu, Hosking and Lenné (2009) used a Honda Rider Trainer (HRT) to compare the effect of having a driving licence on the hazard perception and responding performance of novice and experienced motorcyclists; Shahar, Poulter, Clarke and Crundall (2010) investigated whether riders and drivers could be discriminated on the basis of their performance in responding to hazards; and Di Stasi, Alvarez-Valbuena, Canas, Maldonado, Catena, Antoli and Candido (2009) assessed the mental workload on riders as they were exposed to hazardous situations. Here, as part of a larger project using a number of methods, a HRT was used to make

comparisons of riding performance between returned, continuing and new riders. The aim of the umbrella project was to identify any meaningful differences between these rider types that might be targetted in a new refresher-type course for more experienced riders. It was expected that continuing riders would negotiate the simulated riding tasks with fewer crashes and higher scores than riders belonging to either of the other two groups.

METHOD

Participants

Forty-five licensed motorcycle riders participated – 15 newly licensed, 15 returned, and 15 continuing riders. A new rider was defined as one who held a motorcycle learner permit and had ridden at least 500 km on-road; a continuing rider had held a full motorcycle licence for at least five years and had ridden frequently and for a minimum of 20,000 km during this period; a returned rider had returned to riding within the last six months after having stopped riding for a period of five or more years. The returned rider was required to have ridden at least 20,000 km prior to taking a break from riding and no more than 500 km following their return to riding. All but three of the 45 riders were men. The average age was 33 years (SD=13 years) for the new riders, 54 (SD=9) for the continuing riders, and 53 (SD=11) for the returned riders. While a variety of recruitment methods were used, most of the participants volunteered after hearing about the study by word of mouth and/or through contact with a commercial rider training provider.

Apparatus

Vidotto et al (2008) provide significant detail describing the Honda Rider Trainer (HRT) simulator (see Figure 1). In brief, the motorcycle interface consists of handlebars with throttle and other controls, with brakes and gear change operable by the user's feet, as is the case for a real motorcycle. A virtual ride is played out on a computer monitor. The rider controls their passage through the scenario, interacting with other road users and infrastructure. Steering at all speeds is controlled via the handlebars only as the motorcycle does not lean for cornering.
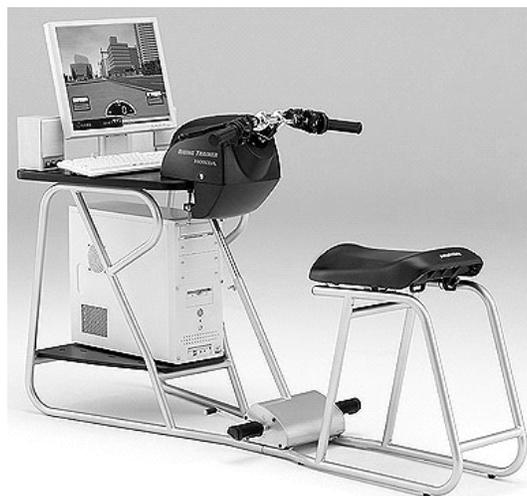
Figure 1. Honda riding simulator

Honda supplies a set of standard scenarios with the simulator. Each scenario is divided into approximately eight segments, or "scenes", and each scene features a particular challenge or hazardous situation to assess the rider's skill and ability. The hazards, such as a child running into the road, are triggered by the rider's location in the scenario. It is possible, and even likely, to crash during a simulation, after which the simulator allows the rider to continue from a stationary start.

In its standard use as a training or assessment tool, the HRT system provides a score from A ('good') to D ('especially beware') for each scene, along with an overall score A-D, on a printed scoresheet. According to Di Stasi et al (2009) A corresponds to "safe behaviour (no accidents, avoided hazards without hard braking or coming too near, followed the speed limits), …[B] = precaution behaviour (didn't follow the speed limit, applied the brakes hard, came near other vehicles)…[C] = hazardous behaviour (hard braking near other vehicles) and …[D] = accident". Each score is accompanied by written feedback of one to several sentences. The scoresheet includes a map of the ride to help the trainee apply the feedback to their memory of the ride.

Procedure

Each participant undertook five rides on the simulator, beginning with a practice session to gain familiarity with the simulator and the task. The practice ride did not include any hazards. The four test rides were presented in counterbalanced order with a break between rides as necessary. Each ride took approximately four to six minutes, depending on the speed of travel selected by the rider and whether they were involved in crashes during the scenario. Two of the test rides were "Path" scenarios and two were "Avenue" scenarios (HRT nomenclature). A Path ride takes place in an urban environment with high traffic density, while an Avenue scenario is a medium traffic density urban environment. Three of the four test rides undertaken for this study consisted of eight scenes and one was composed of seven scenes. According to Vidotto et al (2008), four HRT simulator scenarios are not sufficient to result in boredom and fatigue for the participant.

Participants were told that the test rides may contain hazards to which they might have to respond. They were asked to try to maintain (but not exceed) the posted speed limit, obey all road rules and signs and signals, and avoid being involved in a crash.

RESULTS

Figure 2 displays the distribution of scores (as a percentage) for each rider group for the two types of simulator rides. For both the Path and Avenue sessions riders across all three groups were most likely to earn an A grade, followed by a B then C, and least likely to achieve a D – the lowest score. With the highest percentage of As amongst their scores, continuing riders seem to have performed at the highest level for both types of ride.

The A to D grading scale is ordinal data because the interval between grades is not quantified and may not be equal. To make additional comparisons between the rider groups the letter-scores were arbitrarily converted to numeric scores: A was converted to 4, B to 3, C to 2, and Ds were translated into a score of 1 – thus a higher score represents better performance. Multiplying the distribution of scores by these numeric values provides a non-sophisticated but useful means of

collapsing the scores and making a comparison between the three groups on the basis of a single overall score. For the Path rides this calculation produced a total score of 287 for the new rider group, 312 for the continuing riders, and 294 for the returned riders, indicating that the continuing riders had indeed performed best at this task, followed by the returned riders, and then the new riders. For the Avenue rides, in which there was less traffic, the continuing riders scored 369 overall, with 350 for the new riders, and 339 for the returned riders – continuing riders seem to have performed at the highest level, followed by new and then returned riders.
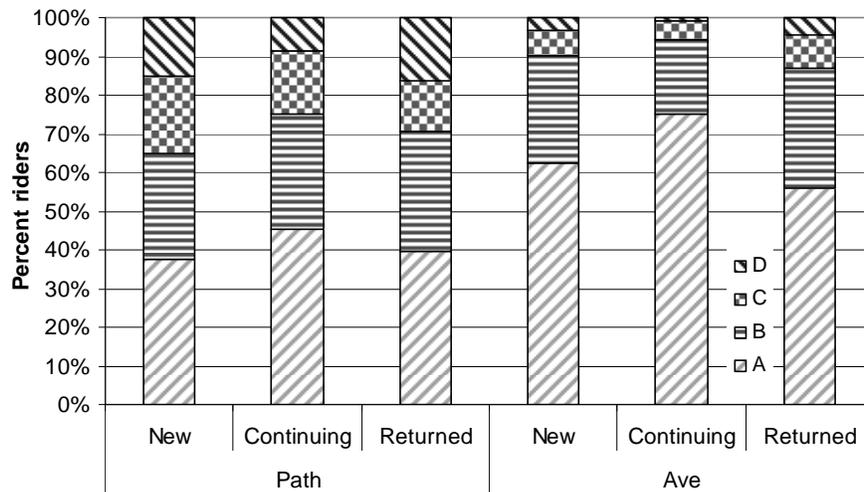
Figure 2. Percentage of each rider group achieving simulator scores A-D for higher-density path and lower-density avenue rides

An appropriate statistical method for making comparisons between more than two sets of ordinal data is the non-parametric Kruskal-Wallis test. With either seven or eight scene scores and one overall score for each scenario – two Path and two Avenue – there are a total of 35 potential comparisons between rider groups. Five of those comparisons were statistically significant at the 0.05 level – all except one were Avenue comparisons. In four instances the continuing riders performed at the highest level, and in four instances the new riders were second-best performers. (The statistics have not been reported here, but are available from the authors.)

Each of the 34 scenes across the four scenarios provided a hazard that could result in a crash. Table 1 summarises the crash profiles as a function of rider type and overall across all riders. At least one rider from each group completed all scenarios without being involved in any crashes. The highest number of crashes for any single rider was 11, for a returned rider. On average, continuing riders were involved in the lowest number of crashes, followed by new, then returned riders. However an AVOVA analysis indicated that these differences were not statistically significant ($F_{(2,42)}=2.6$; $p>0.05$).

Table 1. Summary crash data as a function of rider type

Statistic	New	Continuing	Returned	Overall
Max crashes	8	4	11	11
Mean crashes	2.9	1.4	3.3	2.5
Standard deviation	2.9	1.1	2.8	2.5

DISCUSSION

The use of the Honda training simulator enables all riders to be presented with a consistent scenario that contains a high potential for the rider to be involved in a crash. Riders received 35 scores based on their performance, of which five revealed statistically significant differences between the groups. In four of those comparisons the continuing group achieved the best performance, new riders performed second-best, and returned riders worst. A comparison of crash involvement within the scenarios did not result in an overall significant difference between the groups. On the whole, the data indicate a trend – that continuing riders generally dealt with the simulation task in a superior manner, followed by the new riders and then the returned riders. However, the initial analysis of the data does not reveal a consistent pattern of difference that could provide direction or content for designing a refresher course for returned and other non-novice riders.

The simulation task is somewhat artificial – probably more so as a motorcycle task than for other simulation activities such as car driving. There is little encouragement to suspend disbelief and begin to "feel" like one is riding since the motorcycle does not lean, and thus it does not in any way handle like a real motorcycle, despite the realistic interface involving handlebars, foot pegs, controls, and so on. Liu et al (2010) used the HRT for their investigation of the effect of car driving experience on rider performance, and highlighted similar concerns. They also noted that the use of a standard (flat-panel) monitor for the display may not allow for peripheral vision to shorten hazard detection time – though wearing a helmet while riding also narrows peripheral vision and possibly negates this criticism to some extent. Shahar et al (2010) projected their HRT footage onto a wall, delivering a larger depiction of the road environment, though they noted that at least some of the hazards depicted may not have conveyed the expected degree of "foreshadowing" – advance cues to the pending hazard.

Criticisms of the technology are warranted, but a simulator provides the only environment in which the rider can be intentionally placed "in harm's way" in order to evaluate performance and reactions in an emergency situation. Further, it is critical to note that all participants in this study were subject to the same short-comings and thus none of these considerations can be cited to explain differences in performance between the groups. The task is more concerned with identification of and reaction to hazards to avoid crashes, some road craft and, when being used for training or testing, competent use of the motorcycle controls, and thus high resolution and high fidelity realism may not be critical.

Often when older participants perform more poorly on simulation tasks a common conclusion or qualifier is that the use of such technology may have been more difficult for them to deal with. The new riders in this study were clearly the youngest, while the returned and continuing riders were very similar in age. Given the superior performance of the continuing riders, age cannot be used to explain the poorer performance of the returned riders. As all of the participants were current motorcycle riders and all were exposed to the same simulations in the same manner, it is also unlikely that other factors such as the simulator interface or the realism or fidelity or content of the simulations would explain the differences that did arise between the rider groups.

The HRT is primarily a training and assessment tool, and so its pre-programmed set of scenarios and its standard output of a set of scores ranging from A to D along with commentary and feedback for the rider are probably sufficient and appropriate for that task. Indeed Vidotto et al (2008) found that extended use of the simulator did result in improvement in rider performance in terms of crash reduction and hazard avoidance (though it should be noted that this performance improvement related to performance using the simulator rather than measured performance in terms of riding an actual PTW on the road in real traffic). The lack of flexibility to manipulate the simulator scenarios for the purposes of research is a shortcoming noted by both Liu et al (2009) and Shahar et al (2010). It is also disappointing that the means used by the simulator to determine the A to D scores is not published. These shortcomings prevent a more detailed analysis of rider performance, at least without inventing a new scoring system.

The simulator seems to assign a score of D for any scene in which the rider collides with an introduced hazard, such as a car entering the road immediately in front. Severe braking also seems to result in a lower score, but not a score of D, even if the degree of braking and avoidance is likely to have, in the real world, resulted in the rider probably crashing. Given that the HRT software provides data across a range of variables collected during each ride it would be possible to develop a new (likely) more sophisticated scoring algorithm that results in multiple scores on a number of outcomes. For example, skill and safety could be assessed separately, as they were in Shahar et al's (2010) study using the simulator, though they scored video footage of the simulator ride rather than using the simulator's data.

Liu et al (2009) and Vidotto et al (2008) also used crashes reported by the simulator. However, a closer examination of the simulator outputs (both the in-depth data and the rider score sheets) reveals a possible anomaly. The score sheet includes both a map of the scenario and a scene-by-scene commentary and scoring. On some occasions, particularly when heavy braking takes place, a crash can be noted on the map and in the scene commentary, but not be taken into account in the overall score or be present in the in-depth output. It is not known whether Liu et al and others who analysed the number of crashes (or hazards successfully avoided) used the more accessible rider score sheets or the in-depth data, though the former is probably more likely. Likewise it is not clear how Shahar et al (2010), who counted crashes from video footage rather than using the simulator output, defined a crash. As an experienced rider their video judge might easily have counted highly probable rather than impact crashes (given heavy braking or sudden swerving) as crashes or simply noted the number of occasions in which the simulator was reset due to a crash. It is thus difficult to make comparisons between the current outcomes and those of previous research, though there would seem to be no previous studies in which a simulator was used to assess returned rider performance versus that of continuing and new riders.

CONCLUSION

No other published examples were found in which comparisons were made between any combination of returned, new and continuing riders using a motorcycle simulator like the HRT. In the current study the HRT has revealed that, on balance, continuing riders performed at the highest level, as predicted. Unexpectedly, however, returned riders produced the worst performance, inferior to that of new riders. The lack of a more robust difference between the groups may indicate that the lack of realism delivered by the simulator is insufficient to reveal

those differences. Alternatively, there may *be* no real differences between rider groups, though this conclusion is less likely given other reports and the crash statistics themselves. Other means of comparing these rider types are currently being undertaken by the authors.

ACKNOWLEDGMENT

REFERENCES

BITRE (2010). Road deaths Australia 2009 statistical summary. Canberra, Australia: Bureau of Infrastructure, Transport and Regional Economics.

Di Stasi, L., Alvarez-Valbuena, V., Canas, J., Maldonado, A., Catena, A., Antoli, A., & Candido, A. (2009). Risk behaviour and mental workload: Multimodal assessment techniques applied to motorbike riding simulation. *Transportation Research Part F, 12*(5), 361-370.

Haworth, N., & Mulvihill, C. (2005). *Review of motorcycle training and training.* (Report No. 240). Melbourne, Australia: Monash University Accident Research Centre.

Haworth, N., Mulvihill, C., & Symmons, M. (2002). *Motorcycling After 30.* (Report No.192). Melbourne, Australia: Monash University Accident Research Centre.

Liu, C. C., Hosking, S. G., & Lenné, M. G. (2009). Hazard perception abilities of experienced and novice motorcyclists: An interactive simulator experiment. *Transportation Research Part F, 12*(4), 325-334.

Mulvihill, C., Symmons, M., & Liu, S. (2010). Identify training needs and develop and pilot a training program for returning motorcycle riders. (Confidential Draft Report to VicRoads). Melbourne, Australia: Monash University Accident Research Centre.

Sexton, B., Baughan, C., Elliot, M., & Maycock, G. (2004). *The accident risk of motorcyclists.* (TRL Report 607). Crowthorne, U.K.: Department for Transport.

Shahar, A., Poulter, D., Clarke, D., & Crundall, D. (2010). Motorcyclists' and car drivers' responses to hazards. *Transportation Research Part F, 13*(4), 243-254.

Vidotto, G., Bastianelli, A., Spoto, A., Torre, E., & Sergeys, F. (2008). Using a riding trainer as a tool to improve hazard perception and awareness in teenagers. *Advances in Transportation Studies an International Journal Section B, 16*, 51–60.