ANZIAM J. 63 (EMAC2021) pp.C56–C68, 2022

Modelling implicit pre-cues and collision avoidance in a driving simulator: A pilot study

Ignatius G. McBride[1] Job Fransen[2] Stephen Woodcock[3]
Gustavo de Andrade[4] Andrew Novak[5]
José Angelo Barela[6] Sérgio Tosi Rodrigues[7]

(Received 26 February 2022; revised 8 June 2022)

Abstract

It is well-established that pre-cues, including those observed in an implicit manner, can affect motor skills and reaction times. However, little research currently exists on how pre-cues influence complex motor skills such as driving a car at high speed. This pilot study investigates the effect of implicit pre-cues on collision avoidance under a repeat trial experiment design using a car driving simulator. Seventeen participants (aged 23.8 ± 4.2 years) were included in this investigation, which consisted of four different one-kilometre driving scenarios. This investigation considers two of the four scenarios. Two scenarios had the

DOI:10.21914/anziamj.v63.17160, Published 2022-06-17, as part of the Proceedings of the 15th Biennial Engineering Mathematics and Applications Conference. ISSN 1445-8810.

stimulus of a child crossing the road, however only one of these scenarios had an implicit pre-cue appear before the stimulus. The remaining two scenarios had no stimulus or pre-cue and were included to reduce any learning effect by participants. The proportion of participants who had a collision differed significantly between scenarios with and without a pre-cue. The primary effect size of the pre-cue is modelled using a logistic regression and distributions for point estimators are obtained from bootstrapping results. A power analysis exploring different primary effect sizes is performed to inform sample size considerations for repeat studies. Implications for motor control, such as experiment design and statistical modelling methods, are discussed to inform future large scale trials.

Contents

1 Introduction

It is well known that the reaction time of manual responses to visual targets is affected when advanced information about the approximate target position is provided [3]. Here, we investigate the impact of implicit pre-cues, that is, an incident which a subject may have unknowingly observed [1]. Using a driving simulator, this pilot study provides insights into the data collection and storage aspects for use in a future larger experiment, and also provides feedback on the participants' experience operating the simulator and awareness of the experiment design.

A common approach for pilot studies is to obtain estimates for effect sizes to conduct power analyses to determine sample sizes for definitive repeat studies [7]. The dataset obtained from this pilot is sufficient to perform parametric hypothesis testing on the primary effect of implicit pre-cues on collision avoidance. Primary and secondary effect sizes are explored using generalised linear models [8]. Distributions for point estimators from regressions are obtain by bootstrapping, which give standard errors that are used to construct confidence intervals [9]. Sample size calculations are performed on different effect sizes to inform future repeat studies.

2 Methods

Data were collected from 18 volunteer participants who responded to a survey to take part in a pilot study conducted over one week in November 2017 at the University of São Paulo. The pilot study received ethical approval and all participants signed consent forms that permitted the analysis of data collected from them during the experiment. Data from one participant were excluded from the pilot study, as the subject had previously taken part in a similar study at the same university, which reduced the total number of participants to 17 for the purpose of this pilot study.

The 17 participants (nine male and eight female) had an average age of 23.8 years (standard deviation $= 4.2$ years). The majority of participants listed their occupation as student (ten students and seven others). Ophthalmic issues were reported by ten participants; nine of these participants indicated that they used glasses for vision correction and one reported no vision correction aids. Each participant had a valid driver license and the average length of license held was 5.4 years (standard deviation = 4.0 years). All 17 participants reported that they drove cars, with eight reporting they also drove motorcycles. Participants drove cars on average 4.9 days each week (standard deviation = 2.6) for an average of 34.9 km each day (standard deviation = 31.7 km).

2.1 Procedure

The pilot study consisted of four different 1 km driving scenarios set in a right-hand side drive context, each in an urban environment on a straight four-lane road. The outside lanes had parked cars, with traffic flowing in the middle two lanes. One condition included a pre-cue of a child's feet obstructed by a parked car on the right-hand-side of the road at a set location along the 1 km drive, visible for more than 600 ms and less than 1000 ms, depending on the driver's velocity. The stimulus, which was a child crossing the road, appeared in two of the conditions; one of these conditions included the pre-cue and the other did not. The other two conditions had no pre-cue or stimulus present to limit a learning effect occurring, and for the purpose of this analysis these two conditions can be considered identical treatments.

Immediately prior to taking part in the experiment participants had one 1 km practice drive in the simulator. This simulation had no pre-cue or stimulus and was intended to allow participants to familarise themselves with the features of the driving simulator such as the steering wheel, pedals and monitor. Following the practice drive, participants completed 12 simulations and took an exit survey which included a question about whether they perceived anything which influenced their reactions while driving. The order of the driving simulation scenarios was randomised for each participant. Each participant did three trials of the following scenarios:

- S1 Pre-cue and stimulus of a child crossing the road;
- S2 No pre-cue and stimulus of a child crossing the road;
- S3 No pre-cue and no stimulus of a child crossing the road; and

S4 No pre-cue and no stimulus of a child crossing the road.

Participants were told to drive as they would normally in their daily lives. When the simulation started, the participant's car was parked on the right-hand-side of the road, the participant was told to leave the parking site and drive to the end of the road or until a grey image appeared. Participants were told to drive at 90 km/h in the appropriate right-hand-lane. The road, weather, daytime scenery and traffic conditions were similar for each scenario. The driving simulator captured data for each individual drive on an elapsed driving time scale.

2.2 Measurements

The driving simulator consisted of a 46-inch Samsung TV as the display and a driving cockpit with steering wheel, pedals (accelerator, break and clutch) and a gear stick all from Logitech (model G27). Researchers had a 23-inch Dell monitor to observe the experiment. The driving simulator used STISIM Drive 3.14.01 software on a Dell Precision Workstation T5810 computer. The scenario configuration was set to record data at a frequency of 30 hz, the transmission type was set to automatic, the transmission gears were also changed to create a more realistic effect and the screen resolution for the TV was set to FULL HD (1920×1080p) with the display frequency of 60 hz.

The driving simulator recorded driving time elapsed data for each trial which included: lateral and longitudinal acceleration (m/s^2) of the car; lateral and longitudinal velocity (m/s) of the car; speedometer (km/h); brake pedal and accelerator input counts; steering wheel angle (degrees); lateral position of the stimulus (relative to the midline of the road in metres); minimum time (s) and range (m) to collision with the stimulus. Experiment variables were ordered on an elapsed time scale that started when the participant moved from the parked site until the end of the trail. For each trial, the trial number (1 to 12) was recorded along with the type of trial (S1 to S4), the time and distance that the stimulus appeared in the trial and 'hit stimulus', a binary 0–1 variable which took the value 1 only in the event of a collision.

Table 1: Contingency table of driving simulation scenarios aggregated by the hit stimulus (child crossing the road) dummy variable.

	Scenario				
	S1	S2	S3	S4	Total
Hit stimulus = 0	31	21	51	51	154
Hit stimulus = 1	20	30	0	0	50
Total	51	51	51	51	204

3 Data analysis

Prior to performing any statistical tests, an exploratory data analysis was performed using R (version 4.0.2) in RStudio (version 1.3.1073). The base R statistics package was used to explore the distributions of explanatory variables. In addition, relationships between explanatory variables were explored using contingency tables for categorical variables and correlations for continuous variables. Categorical experiment data were explored using frequency tables and the relationships between individual numerical variables with a correlation matrix. Various contingency tables were constructed using the 'hit stimulus' variable to investigate the relationships with other variables captured from the driving simulator summaries. Table 1 present the frequency of each scenario and stimulus.

Scenarios S3 and S4 did not include a stimulus in the simulation, and accordingly there are no hit stimulus results in Table 1. Exploring the correlations between the numeric explanatory variables in scenarios S1 and S2 displayed evidence of multicollinearity. For example, the explanatory variables 'time child hit' and 'elapsed driving distance' were only populated in the summary extracts when the child was hit, making the information redundant for modelling using 'hit stimulus' as a response variable.

3.1 Data methods

The scenario proportions, calculated from the binomial 'hit stimulus' variable, are the statistics for comparison in a t-test. A significant difference between two proportions gives evidence that the pre-cue in S1 influences the response of drivers to the child stimulus present in experiment scenarios S1 and S2. With p_{S1} and p_{S2} defined, respectively, as the proportion of simulations with collisions in scenarios S1 and S2. We tested two hypotheses:

$$\begin{aligned} H0: &\quad p_{S1} - p_{S2} = 0\,, \\ H1: &\quad p_{S1} - p_{S2} \neq 0\,. \end{aligned}$$

The two proportion t-test was performed in R using the `prop.test()` function, giving a p-value and confidence interval for the true difference in proportions. The significance level was set using the `conf.level` argument within the `prop.test()` function. In this instance using a significance level of 0.05, entered as a `conf.level` of 0.95.

A power analysis was performed using the 'pwr' package. Holding the scenario S2 proportion constant at 0.5 and using the `pwr.2p.test()` function, a number of scenario S1 effect sizes between 0.3 and 0.5 were explored to obtain sample sizes for a two-sided alternative hypothesis. This process was repeated for powers 0.4, 0.5, 0.6, 0.7 and 0.8, and a constant significance level $(\alpha = 0.05)$ to determine sample size requirements for future studies.

Generalised linear models were used to fit to non-normal response variables, in this case the binomial 'hit stimulus' variable. The expected value of the response is related to the linear predictor using a link function. A logit link function, also called log-odds, is

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = \beta_0 + \beta_1(\text{scenario})\,.$$

Converting the response variable to odds, the probability of hitting the stimulus, calculated as the sum of stimulus hits divided by the number of trials, is $\hat{p} = (20 + 30)/(51 + 51) = 0.4902$. The intercept parameter is β_0, and β_1 is the parameter for the scenario factor variable. The `glm()` function was used to fit logistic regressions with a logit link function. Fitted model estimates were bootstrapped in R, by resampling with replacement from the pilot data 10 000 times, to determine distributions for the intercept and scenario parameters, and Figure 1 plots these two distributions.

4 Results

A two-side proportion t-test was performed, using a significance level of 5% without continuity correction, which gave a Pearson Chi-squared value of $\chi^2 = 3.9231$ (with $df = 1$) and determined a significant difference (p-value = 0.0476) between the two proportions. We conclude that the presence of a pre-cue in scenario S1 is associated with a reduction in the proportion of drivers hitting the stimulus compared to scenario S2. The corresponding 95% confidence interval for the difference in proportions was $[-0.3863, -0.0058]$.

Taking the arcsine square root transformation of the observed difference in proportions between scenario S1 and scenario S2 gave an effect size of $h = 0.3948$, and using a significance level of $\alpha = 0.05$ and power $1 - \beta = 0.8$ resulted in a minimum sample size requirement of 101 [2]. Holding the power and significance level constant, a smaller effect size of $h = 0.2948$ would require a minimum sample size of 181 to determine a significant difference between scenario S1 and scenario S2. Figure 2 plots the required sample size for different effect sizes and powers.

A logistic regression was performed on the response variable 'hit stimulus' and independent factor variable 'scenario'. Fitting the model gave an intercept parameter $\beta_0 = -0.4383$, with standard error $= 0.2868$ and scenario parameter $\beta_1 = 0.7949$, with standard error $= 0.4040$. The estimated 'scenario' parameter is statistically significant (p-value $= 0.0491$), and taking its exponential gives an odds ratio of 2.2143 which indicates the odds of hitting the stimulus in scenario S2 are 2.21 times higher than in scenario S1.

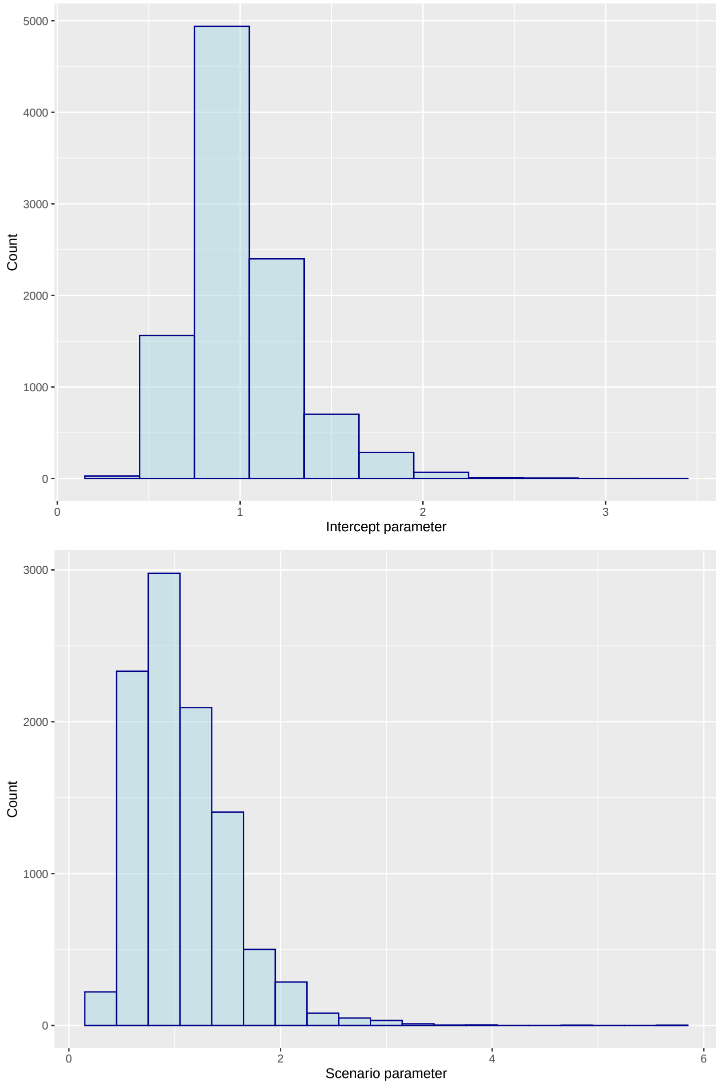

Figure 1: Bootstrapping yielded distributions for the intercept parameter (top) and scenario parameter (bottom) in the logistic regression. The scenario parameter distribution is unimodal and positively skewed, indicating that no pre-cue leads to more collisions.

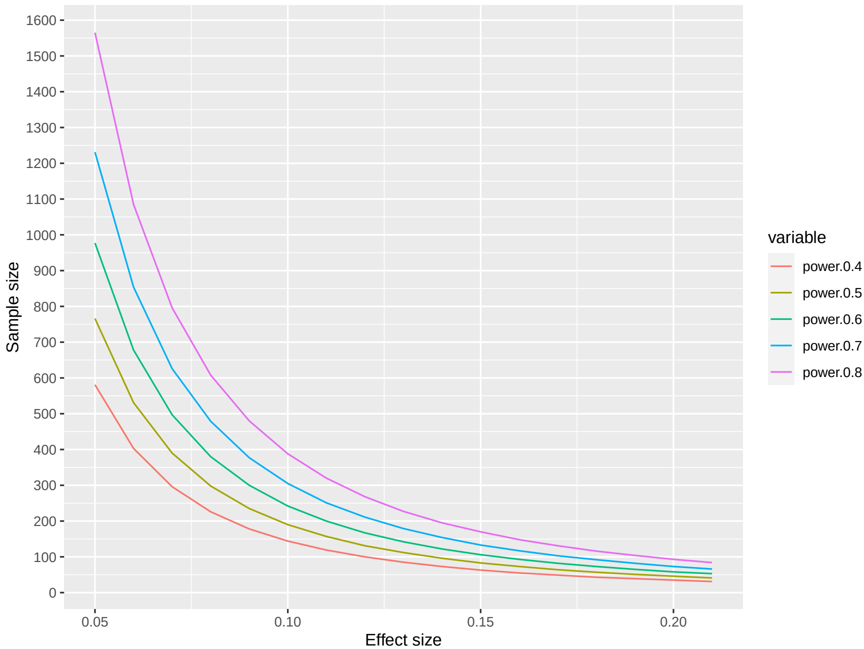

Figure 2: Smaller effects require larger sample sizes. Similarly, holding the effect size constant, the higher the power the larger the sample size required to obtain statistically significant results in a repeat experiment.

5 Discussion

Pilot studies present researchers with the opportunity to field test logistical aspects of experiment design and obtain preliminary answers to primary research questions. Glitches in software or measurement devices can be addressed during pilot studies [6]. Pilot studies are also an important mechanism to train researchers in experiment protocols as well as gain insight into the availability of willing participants and recruitment design. This pilot study provided internal validity of the primary effect of an implicit pre-cue on collision avoidance and the generalisability of results using different modelling approaches in the field of skill acquisition.

Examining complex motor responses to a stimulus, with the presence of an implicit pre-cue, provides insights into movement preparation and road safety. We found a significant difference between the proportion of collisions in the scenario with an implicit pre-cue appearing before the stimulus compared the scenario without this pre-cue. It would be reasonable to expect similar findings in repeat studies. Interestingly, post-experiment, no participant knowingly identified the presence of a pre-cue in the experiment. This observation may inform the design phase of future analyses. The post experiment survey allowed for participants to be grouped into blocks, based on, for example, corrective vision aids and years of driving experience, for further explanatory modelling. No statistically significant model parameters were obtained from exploring various blocks using generalised linear mixed effect models [5].

Data from this pilot study may help inform sample size calculations for future analyses. Based on other studies of model diagnostics, samples of at least 20 are suggested for ordinary least squared regressions [4]. The number of trials in the pilot study is sufficient to explore simple logistic models, however more trials are needed to investigate multivariate mixed effects models recognising there are several goodness-of-fit tests for logistic regressions. Model diagnostics and validation methods such as split-sample and cross validation should be considered for future studies which employ classification models to explore secondary effects—such as vision impairment or driving experience.

The bootstrapped distributions of the intercept and scenario parameters exhibited characteristics of approximately normal distributions. Under a Bayesian framework, these distributions could be used as informed priors to perform Bayesian logistic regressions. This framework also provides credible intervals for model parameter estimates, rather than confidence intervals, which is a practical alternative to communicate the uncertainty of fitted parameters to health and sport science practitioners. Any repeat experiment should contain a scenario with a pre-cue and no stimulus, which would give further insights into complex motor responses to implicit pre-cues.

References

[1] J. A. Barela, A. A. Rocha, A. R. Novak, J. Fransen, and G. A. Figueiredo. "Age differences in the use of implicit visual cues in a response time task". In: *Braz. J. Motor Behav.* 13.2 (2019), pp. 86–93. DOI: 10.20338/bjmb.v13i2.139 (cit. on p. C58).

[2] J. Cohen. *Statistical power analysis for the behavioral sciences.* Routledge, 1988. DOI: 10.4324/9780203771587 (cit. on p. C63).

[3] U. Eversheim and O. Bock. "The role of precues in the preparation of motor responses in humans". In: *J. Mot. Behav.* 34.3 (2002), pp. 271–276. DOI: 10.1080/00222890209601945 (cit. on p. C57).

[4] D. G. Jenkins and P. F. Quintana-Ascencio. "A solution to minimum sample size for regressions". In: *PLOS One* 15.2 (2020), e0229345. DOI: 10.1371/journal.pone.0229345 (cit. on p. C66).

[5] J. Jiang. *Linear and generalized linear mixed models and their applications.* Springer Series in Statistics. Springer, 2007. DOI: 10.1007/978-0-387-47946-0 (cit. on p. C66).

[6] C. Kistin and M. Silverstein. "Pilot studies: A critical but potentially misused component of interventional research". In: *JAMA* 314.15 (2015), pp. 1561–1562. DOI: 10.1001/jama.2015.10962 (cit. on p. C65).

[7] H. C. Kraemer, J. Mintz, A. Noda, J. Tinklenberg, and J. A. Yesavage. "Caution regarding the use of pilot studies to guide power calculations for study proposals". In: *Arch. Gen. Psych.* 63.5 (2006), pp. 484–489. DOI: 10.1001/archpsyc.63.5.484 (cit. on p. C58).

[8] J. A. Nelder and R. W. M. Wedderburn. "Generalized linear models". In: *J. Roy. Stat. Soc.* 135.3 (1972), pp. 370–384. DOI: 10.2307/2344614 (cit. on p. C58).

[9] R. Stine. "An introduction to bootstrap methods: Examples and ideas". In: *Soc. Meth. Res.* 18.2–3 (1989), pp. 243–291. DOI: 10.1177/0049124189018002003 (cit. on p. C58).

Author addresses

1. **Ignatius G. McBride**, School of Mathematical and Physical Sciences, University of Technology, Sydney 2007, Australia.
   mailto:Ignatius.G.McBride@student.uts.edu.au
   orcid:0000-0001-5394-1785
2. **Job Fransen**, Human Performance Research Centre, University of Technology, Sydney 2007, Australia.
   mailto:job.fransen@uts.edu.au
3. **Stephen Woodcock**, School of Mathematical and Physical Sciences, University of Technology, Sydney 2007, Australia.
   mailto:stephen.woodcock@uts.edu.au
4. **Gustavo de Andrade**, Laboratory of Information, Vision and Action, Department of Physical Education, Faculty of Sciences, São Paulo State University, Brazil.
   mailto:gustavo.andrade@unesp.br
5. **Andrew Novak**, Human Performance Research Centre, University of Technology, Sydney 2007, Australia.
   mailto:andrew.novak@uts.edu.au
6. **José Angelo Barela**, Department of Physical Education, Institute of Biosciences, São Paulo State University, Rio Claro, SP, Brazil.
   mailto:jose.barela@unesp.br
7. **Sérgio Tosi Rodrigues**, Laboratory of Information, Vision and Action, Department of Physical Education, Faculty of Sciences, São Paulo State University, Brazil.
   mailto:sergio.tosi@unesp.br