EFFECTS OF HANDS-FREE CELLULAR PHONE CONVERSATIONAL COGNITIVE TASKS ON DRIVING STABILITY BASED ON DRIVING SIMULATION EXPERIMENT

Wei YAN[a], Wang XIANG[b]*, S.C. WONG[a], Xuedong YAN[c], Y.C. LI[a], Wei HAO*[b]

[a] *Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, China*

[b] *Key Laboratory of Special Environment Road Engineering of Hunan Province，Changsha University of Science & Technology, China*

[c] *Key Laboratory of Ministry of Education – Urban Transportation Complex Systems Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, China*

** Corresponding author: Tel.: +86 15116988017, E-mail: 158995646@qq.com,*

or Tel.: +86 15343318289, E-mail: haowei@csust.edu.cn

ABSTRACT

Driver distraction due to cellular phone usage is a major contributing factor to road crashes. This study compares the effects of conversational cognitive tasks using hands-free cellular phone on driving performance under three distraction conditions: (1) no distraction (no cellular conversation), (2) normal conversation (non-emotional cellular conversation), and (3) seven-level mathematical calculations. A car-following scenario was implemented using a driving simulator. Thirty young drivers with an average age of 24.1 years maintained a constant speed and distance between the subject vehicle and a leading vehicle on the driving simulator, and then respond to the leading vehicle's emergency stop. The driving performances were assessed by collecting and statistically analyzing several variables of maneuver stability: the drivers' brake reaction times, driving speed fluctuation, car-following distance undulation, and car-following time-headway undulation. The results revealed that normal conversation on a hands-free cellular phone impaired driving performance. The degree of impairment caused by normal calculation was equivalent to the distraction caused by Level 3 mathematical calculations according to the seven-level calculation baseline. The calculation difficulty of Level 3 is one double-digit figure plus a single-digit figure, and non-carry addition mental arithmetic is required, e.g., 44 + 4. The results indicated that an increase in the level of complexity of the calculation task was associated with an increase in brake reaction time. The seven-level calculation-task baseline could be applied to measure additional distraction effects on driving performance for further comparison.

Keywords: Driving Simulator, Hands-free Cellular Phone Conversational Cognitive Task, Road Safety, Driving Performance Stability, Mathematical Calculation

1 INTRODUCTION

Being distracted while driving is considered one of the most common and severe causes of traffic crashes. Distracted driving caused by cellular phone usage is a growing and severe threat to road safety (WHO, 2011). Researchers have found that the risk and driving performance impairment from talking on either handheld or hands-free cellular phones (Fitch et al., 2015, Fitch et al. 2013, Patten et al., 2004, Hancock et al., 2003, Yan et al., 2015); while driving impairs drivers' physical performance in maneuvering their vehicles, takes their eyes off the road, and increases their mental workload by taking their minds off the driving situation (Ishigami and Klein, 2009, Harbluk et al., 2007, Caird et al., 2008, Li et al., 2016); and the impact of hands-free mobile phone does not provide greater safety as compared to hand-held mobile phones while driving (Caird et al., 2008, Lipovac, 2017). The impaired performance includes a longer brake reaction time, an undulating lane course, a fluctuating driving speed, an inconsistent following distance and time headway, and a failure to remember having seen objects (Atchley and Dressel, 2004, Caird et al., 2008, Drews et al., 2008, Horrey and Wickens, 2006, Rakauskas et al., 2004, Strayer and Johnston, 2001, Strayer et al., 2003, Rosenbloom, 2006, Yan et al., 2015). On the contrary, some other recent studies have failed to find strong links between conversations using hands-free devices and the risk of a safety-critical event (NHTSA, 2016, Fitch et al., 2013, Simmons et al., 2016, Oviedo-Trespalacios et al., 2016, Cunningham et al., 2017).

Many jurisdictions have prohibited handheld cellular phone use while driving. In Hong Kong SAR, the use of handheld cellular phones during driving has been banned since July 1, 2000 (Road Safety Council, Hong Kong, 2003), however, hands-free cellular phone use is still prevalent in Hong Kong. Continuously increasing number of states in the US have banned cellular phone usage while driving, as well as many cities in China and the United Kingdom, handheld cellular phone usage while driving incurs penalty points added to the driving license and an additional fine. However, the law does not specifically prohibit or control the types of hands-free accessories that drivers use for their cellular phones. Because the risk of crashing and the impairment of driving performance caused by using hands-free cellular phones are not commonly agreed and well recognized by the public. Some jurisdictions have enacted stricter laws to ban the use of hands-free devices, such as Japan. Whether these jurisdictions should further strengthen these laws has become a prevalent topic for road safety. Young drivers are the group most commonly impaired by cellular phone usage while driving (Strayer and Drews, 2004, Lipovac et al., 2017, Trivedi et al., 2017).

For driving performance impairment caused by hands-free cellular phone, driving simulator studies, naturalistic driving studies, and combined studies have revealed that the average brake reaction time (BRT) increases when drivers talk on cellular phones as the most prevalent indicator for driving performance evaluation; the lateral control impaired with increased driving lane undulation (DLU), although lateral position control is not a significant variable for some performance impairment studies, because the ability to control lateral position becomes unstable only if the drivers are involved in difficult tasks; and the ability to maintain a consistent driving speed and longitudinal speed control is significantly impaired as well. The driving headway, car-following distance undulation (CDU), and time headway caused by hands-free cellular conversation impairment. Therefore, a driver's BRT, driving speed fluctuation, CDU, and time headway undulation are the most effective indicators of impaired stability performance caused by hands-free cellular phone conversation distraction, the literatures are listed in Table 1.

Table 1. Driving-performance - dependent sample variables, descriptions, and references for cellular phone usage studies.

Variable Classification	Variable	Description	Sample Reference
Brake reaction time	Brake reaction time (BRT)	Time from seeing a hazard to the onset of brake application	Caird et al. (2008); Horrey and Wickens (2006); Al-Darrab et al. (2009); Lamble et al. (1999); Charlton (2009), Ålm and Nilsson, (1995); Strayer and Johnston (2001); Strayer et al. (2003); Caird et al. (2014).
Lateral control	Driving lane undulation (DLU) or SD of lane position (SDLP)	SD of the lateral position	Ålm and Nilsson (1994); Rakauskas et al. (2004); Beede and Kass (2006); Brookhuis et al. (1991); Liu and Lee (2006); Shinar et al. (2005); Caird et al. (2014).
Longitudinal control	Speed of driving speed, speed fluctuation (DSF)	SD of speed	Liu and Lee (2005); Cooper et al. (2003); Ålm and Nilsson (1994); Beede and Kass (2006); Rakauskas et al. (2004); Shinar et al. (2005); Strayer et al. (2003); Fitch et al. (2013); Thapa et al. (2015).
	Car-following distance, distance undulation (CDU) Car-following time headway undulation (CTU)	SD of the car-following distance to the rear bumper of the lead vehicle	Strayer et al. (2003); Strayer and Drews (2004); Ålm and Nilsson (1994); Ålm and Nilsson (1995); Caird et al. (2014).

Many studies have measured the impairments caused by cellular phone conversations. Some used verbal recall or recognition tasks (Haigney et al., 2000, Mazzae et al., 2004, Strayer and Johnston, 2001), and others used naturalistic phone conversations to measure the effects of cellular phone usage (Rakauskas et al., 2004, Shinar et al., 2005). However, the levels of impairment in such studies cannot be measured and compared. To standardize the distraction levels, some studies have incorporated mathematical calculations (Brookhuis et al., 1991, McKnight and McKnight, 1993, Patten et al., 2004). Mathematical calculations were first used in distraction analysis by McKnight and McKnight (1993), and holding a complex phone conversation was set as the distraction to solving a mathematical problem. This distraction condition, and those of dialing a cellular phone, holding a simple phone conversation, and tuning a radio, were compared with the baseline condition of no distraction. Patten et al. (2004) compared both a simple conversation of repeating back single digits and the solving of arithmetic problems with the baseline condition of no distraction to understand the effects of different distraction tasks on driving performance. Harbluk et al. (2002) further studied the effect of mathematical calculations by specifying two calculation difficulty levels: easy addition problems (e.g., 6 + 9) and complex addition tasks (e.g., 47 + 38). These studies found that drivers' perceptions of workload increased and that their driving safety level decreased with the increasing level of calculation. It is evident that mathematical calculations can be used as a standard benchmark to measure distraction impairment. Hence, the accurate standardization of the measurement of driving impairment caused by cellular phone conversation is a worthwhile aim, because sorting of calculation tasks by level allows refined mathematical calculations of the effects of conversations.

Methods based on driving simulators have been frequently used to evaluate the performance of drivers distracted by cellular phone usage and to examine the effects of conversation (Törnros and Bolling, 2005, Beede and Kass, 2006, Ålm and Nilsson, 1995, Drews et al., 2008, Saifuzzaman et al., 2015, Horrey and Wickens, 2006, Rakauskas et al., 2004, Ålm and Nilsson, 1994, Maciej et al., 2011). In this study, a driving simulator was used to model the effects of hands-free cellular phone conversation on the driving performance of young drivers, compared with benchmarked distraction tasks involving calculations. Data were collected to measure driving performance under various distraction conditions; several performance measures, including BRT, DSF, CDU, and CTU, were compared; and the distraction levels were measured. Demographic effects and the mechanism of distraction were also studied. Moreover, a standardized distraction measurement system was developed.

2 METHODS

2.1 Participants

Thirty young Chinese drivers (19 males and 11 female) between 22 and 33 years of age (mean, 24.1 years; SD, 2.4 years) were recruited from a university for this driving simulation study. Those who felt dizzy or other uncomfortable were excluded at the trial stage. All of the participants had valid full driving licenses, and they had had them for periods ranging from 6 months to 11 years (mean = 3.6, SD = 2.3). Most participants were occasional drivers who drove an average of 2.2 hours per week (SD, 3.0 years). Among the participants, 20% had engaged in cellular phone conversations while driving. Half of all drivers had no cellular phone-using habit. All of the participants were Mandarin native speakers. Each participant was invited to attend two experimental sessions: a trial session for familiarization with the driving simulator and the test session. A driving behavior questionnaire was collected by each participant before the experiment. A souvenir worth HK$50 was used to reward each driver’s participation.

The Human Research Ethics Committee for Non-Clinical Faculties of The University of Hong Kong approved the simulated driving experiment. The purpose and experimental procedures of the study were explained and clarified to the participants, and informed consent was obtained before the simulation study was conducted. All participants were fully aware of the experiment’s purpose, procedures, and potential risks beforehand.

2.2 Subjects/apparatus

A desktop-based driving simulator (XPDS 300 Driving Simulator, version 1.6) comprising a driving scenario engine, three 19″ LCD monitors, and a Logitech G27 steering wheel and foot pedal control kit was used to study the effects of hands-free cellular phone conversational cognitive tasks. Data on driving performance stability, including vehicle position, travel speed, acceleration and braking performance, were recorded in a 30-Hz sampling frame and statistically analyzed.

2.3 Driving scenarios

The Free Drive scenario was applied in the trial session and the beginning of the test session to screen the participants and enable them to practice maneuvering the driving simulator. At the beginning of the experiments, the participants were required to familiarize themselves with the driving simulator by driving in the left lane of the motorway at a speed of 50 km/h, to ensure similarity with real driving conditions. This scenario prepared the participants for the driving simulation system, and to minimize the maneuvering differences between the simulation and real driving conditions.

The 2-s car-following scenario was applied in the test sessions. This scenario was conducted to monitor and measure the effects of conversational tasks on driving stability and to compare with the effects of conversational tasks using hands-free cellular phone when driving on an urban road. During the driving tasks, the participants were required to appropriately follow the leading vehicle. During each test run, the leading vehicle accelerated to the prescribed speed of 50 km/h, which it then maintained. The driver was then asked to accelerate to the prescribed speed of 50 km/h and to maintain that speed along an urban road by following the leading vehicle at a safe distance as per the 2-s rule. When the leading vehicle began to brake (indicated by its rear brake light), the participant was required to respond to the brake and come to a complete stop. Figure 1 illustrates the testing scenario of the driving simulator.

Figure 1. Testing scenario of the XPDS 300 Driving Simulator.

In the test process, the driver's performance stability was assessed by his or her ability to maintain the prescribed speed and keep a safe distance from the leading vehicle along a straight road section, along with the BRT in response to the leading vehicle's braking. The BRT of the leading vehicle was randomly generated to avoid interference from the learning effect. Each condition was simulated twice for data extraction, and the average of the two runs was applied to increase the reliability of the data.

2.4 Cellular phone

During the hands-free cellular phone's conversational experiment, a Samsung Galaxy S Android smartphone was provided to the drivers together with a set of wireless Bluetooth headphones.

Before the experimental sessions, the participants were allowed to familiarize themselves with the headphones, adjust the volume, and practice using them while driving. Only hands-free conversations were involved in this experiment; no dialing or screen-touching was required. Before each test run, the phone call was pre-connected so that the conversation was already in progress. The participants then followed the instruction to respond to conversational questions or mathematical calculation tasks, according to the simulated driving scenario.

2.5 Distraction Tasks

The experiment involved two types of distraction: mathematical calculations with various levels of difficulty and a set of non-emotional questions typical of normal conversation. An invigilator asked the mathematical problem or made normal conversation from a soundproofed room away from the driving simulator site. During each test run which was under the distraction conditions, the participant was asked to talk continuously on the cellular phone throughout the driving task to ensure continuous and uninterrupted distraction (from the time they started the test car until the onset of the leading car's rear brake light), which marked the end of each test.

The calculation tasks were designed with seven levels of mathematical difficulty, which induced various levels of mental workload in the distracted drivers. In the driving simulator experiment, the order of the difficulty levels of the mathematical calculations was randomized by use of the counterbalance effect. Table 2 shows the specific design criteria and examples of mathematical calculations with various levels of difficulty.

Table 2. Mathematical calculation difficulty: standards and examples.

	Calculation type (number of digits in addends)		Calculation of decimal carry-unit digit	Calculation of decimal carry-10 digit	Examples	
Level 1	Output less than 20	1-digit + 1-digit	N	N	1+6 2+4	1+7 2+5
Level 2		1-digit + 1-digit	Y	N	8+9 6+8	9+9 7+8
Level 3	Output less than 100	1-digit + 2-digit	N	N	31+6 44+4	33+5 52+7
Level 4		1-digit + 2-digit	Y	N	42+9 66+5	56+6 69+3
Level 5	Output less than 200	2-digit + 2-digit	N	N	48+51 64+24	54+34 71+16
Level 6		2-digit + 2-digit	Y/N	N/Y	35+46 57+28	39+80 47+72
Level 7		2-digit + 2-digit	Y	Y	56+67 99+88	99+99 77+88

For the calculations, the participants were asked to perform the addition of positive integers. The seven levels of difficulty were defined based on three criteria: (1) the result of the addition, (2) the number of digits in the addends, and (3) the calculation (or not) of a decimal carry-unit digit. The seven levels of difficulty were designed to create a measurable, continuous benchmark for evaluating distraction levels. The validity and rationality of the difficulty levels were checked according to the average calculation time and accuracy of each level were collected and processed. In each test, data was recorded on the subject's reactions and the number of calculations completed. The calculation time and calculation mistakes increased linearly as the level of difficulty increased.

"Normal" conversation is defined as conversation on personal everyday topics, such as food, weather, or pets, not involving any emotional discussion or any topics with privacy concerns. Samples of the normal conversation questions are listed in the Appendix. Normal conversations were conducted for comparison analysis with the benchmark of the mathematical addition calculations.

2.6 Experimental procedure

The driving simulator experiment consisted of four phases, all experimental sessions lasted around 1 hour, as illustrated in Figure 4.

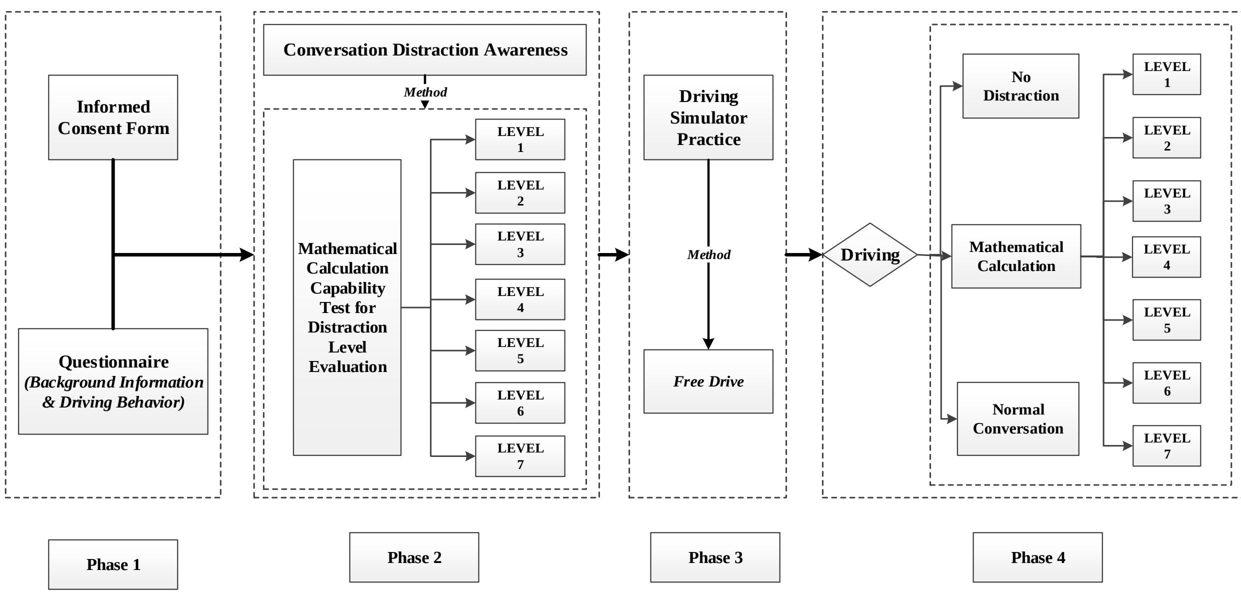


Figure 2. Procedures for cellular phone distraction experiment.

In Phase 1, an invigilator explained the purpose and experimental procedures to the participants, an informed consent form and a driver behavior questionnaire including participants' demographic information were collected.

In Phase 2, each participant took a test of his or her mathematical calculation capability by conversational calculation task using a Bluetooth headphone to collect the baseline of each participant's calculation-time data for the seven calculation levels, and to record the number of calculation mistakes for each level.

In Phase 3, the participants took part in 20-minute training session by using Free Drive scenario to familiarize themselves with the driving simulator. The participants were then involved in normal conversations and answering sample mathematical questions, delivered via the Bluetooth headphone while driving, for the purpose of equipment adjustment.

In Phase 4, the participants were asked to complete a series of simulated driving tests under different distraction conditions: (i) no distraction, (ii) normal conversation, and (iii) seven-level mathematical calculation. In each single run, all mathematical calculations were set at the same difficulty level but were asked in continuously in a random order. The contents of the normal questions were chosen randomly based on a prepared question bank.

3 Data and Analysis Methodology

To study the distraction caused by hands-free cellular phone conversation, we used analysis of variance (ANOVA) and K-S/median test comparison to assess the impairment of driving performance by normal conversation and by each calculation level. This study developed benchmarks for standardized distraction levels using indicators including BRT, DSF, CDU, and CTU to measure the impairment of driving performance by over-the-phone conversations. A structural equation model was then applied to analyze the mechanism of the effects.

3.1 Explanatory variables

The participants' driving performances were measured by BRT, DSF, and CDU under different distraction conditions, including no distraction, normal conversation, and mathematical calculation. In this study, as each experiment was performed twice, each indicator was taken as the average value of the two experiment tests.

(a) Brake reaction time (BRT)

BRT reflects the time interval between the appearance of the leading vehicle's brake light and the time at which the test vehicle driver applied the brake. Hence, it includes the participant's perception and action times.

(b) Driving speed fluctuation (DSF)

DSF measures the fluctuation in a test vehicle's speed and represents a driver's ability to follow a leading vehicle steadily.

(c) Car-following distance undulation (CDU)

CDU measures the fluctuating distance between a leading vehicle and a test vehicle and represents the test vehicle driver's ability to follow the leading vehicle, driving preference, and response to the leading vehicle's performance.

(d) Car-following time headway undulation (CTU)

CTU measures the fluctuation in time headway between a leading car and a test car and represents the test car driver's ability to follow the leading car, based on his or her driving preference and response to the leading car's time headway measurement.

Table 3 presents a summary of the data. The average BRT was 0.85 s with an SD of 0.22 s. The DSF was 1.02 km/h (SD, 0.44 km/h) when the expected driving speed was 50 km/h. The average CDU was 3.55 m (SD, 1.76 m), based on the 2-s rule for distances further than 27.78 m at a 50 km/h driving speed. The average CTU was 0.32 s (SD, 0.14 s), giving a car-following time headway distance of 2 s, with a 16% undulation range.

Table 3. Summary of statistics of the driving performance explanatories of stability.

	Mean	SD	Min.	Max.
BRT (s)	0.85	0.22	0.50	2.00
DSF (km/h)	1.07	0.44	0.10	2.70
CDU (m)	3.55	1.76	0.76	13.10
CTU (s)	0.32	0.14	0.30	1.05

BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.

To evaluate the impairment effects of distraction on driving performance under the normal conversation condition and mathematical calculation, as compared with the baseline condition of no distraction (in which no conversation was made), ANOVA was used to analyze the differences between the group means of the driving performance indicators of BRT, DSF, CDU, and CTU.

3.2 Demographic effects

To analyze the effect of demographic factors on driving performance, we first denoted the four measurement variables, BRT, DSF, CDU, and CTU, as the k of participant i, and then divided the set of driving performance factors f_{ki} into two groups according to the demographic factor being assessed, i.e., $\left(f_{ki}^{1}, \forall i \in G_1\right)$ and $\left(f_{ki}^{2}, \forall i \in G_2\right)$, where G_1 and G_2 are two demographic groups. $G_1 \cap G_2 = \varnothing$, i.e., an empty set, and $G_1 \cup G_2 = M$, i.e., a set comprising all participants. We examined the effects of participants' gender (female = 0, male = 1), age (25 or above = 1 versus below 25 = 0), number of years with a full driving license (more than 3 years = 1 versus 3 years or less = 0), driving frequency (more than once per week = 1 versus once or less than once per week = 0), driving duration per week (more than 3 hours = 1 versus 3 hours or less = 0), and whether the participants had ever talked on cellular phones while driving (yes = 1 versus no = 0). A standard ANOVA analysis of each pair of demographic factor data was conducted to ascertain the effect of demographic factors on driving performance factors. Because of the small sample size, the ANOVA tests are all at the 1% level of significance.

3.3 Comparison between the calculation and normal conversation conditions

A non-parametric Kruskal-Wallis test was conducted to directly compare the distributions of the conversation levels using differential analysis of various surface parameters. The Kruskal-Wallis test is a form of ANOVA performed on ranks. The test assesses the hypotheses that the samples being compared are drawn from the same distribution or from distributions with the same median. A chi-square test was used to test for the differences between two categorical variables. Differences were considered statistically significant for p values of less than 0.05. The Kruskal-

Wallis test is very sensitive to shifts in distribution if all of the k distribution density functions have a similar form; it is based on the ranks of observed values instead of the values themselves (Sunyaev et al., 1998). First, the initial data from all of the k samples are sorted. R_{ij} is the rank of the ith observation in the jth sample. The Kruskal-Wallis statistic is expressed as follows:

$$K = \left(\frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{n_j}\right) - 3(N+1) \text{ and}$$

$$R_j = \sum_{i=1}^{n_j} R_{ij} \text{ ,}$$

where N is the total number of observations in all of the samples, n_j is the number of values in the jth sample, and R_j is the sum of the ranks in the jth sample.

The Kruskal-Wallis non-parametric test is used to compare the distributions, and the median test is used to verify the similarity of the distributions. The null hypothesis in both cases is that the values are identical in all groups. The median test combines each group of test data on at least an ordinal scale. It tests the null hypothesis that the medians of the populations from which two or more samples are drawn are identical. In this study, the driving performance data for each calculation level were assigned to two groups, one consisting of values higher than the median and the other consisting of values at or below the median. Pearson's chi-square test was used to determine whether the observed frequencies in each sample differed from the expected frequencies derived from the distribution of the combined groups. The similarity was weighted by the ratio of frequencies.

The median test was used to examine the frequency ratio of BRT, DSF, CDU, and CTU for each calculation condition. The frequency ratio difference of each driving performance factor was calculated for each calculation condition. The absolute differences of the frequency ratios between the normal conversation and the seven calculation conditions were then calculated; these data were used in the median test to assess the similarity in distribution between the normal conversation condition and each difficulty level of the calculation task condition.

The distributions of the driving performance factors were not consistent; BRT was the principal factor in the distraction experiment, and DSF, CDU, and CTU all measured driving maneuver stability. Thus, the calculation weight for the model used in comprehensive analysis of the driving performance factors was set as 3:1:1:1 for BRT: DSF: CDU: CTU. This analysis was used to evaluate the differences in the level of distraction impairment between the NC condition and each calculation task condition. The null hypothesis held that both groups were sampled from populations with identical distributions. The smaller the difference between two distributions, the higher the similarity of normal conversation to the tested calculation task level.

4 ANALYSIS AND RESULTS

4.1 Normal conversation versus no distraction

The results of the ANOVA comparison of driving performance factors under normal conversation (NC) and no distraction (ND) conditions are given in Table 4. The drivers' BRT under the normal conversation condition was clearly slower than that under the no distraction condition, with an absolute difference of 0.156 s and a relative difference of 21.5%. DSF under the normal conversation condition was more severe than under the no distraction condition, with an absolute difference of 0.245 km/h and a relative difference of 29.8%. Comparing the means, CDU under the normal conversation condition was less stable than under the no distraction condition, with an absolute difference of 0.207 m and a relative difference of 7.0%; CTU under the normal conversation condition was more unstable than under the no distraction condition, with an absolute difference of 0.039 s and a relative difference of 14.9%. However, the differences in CDU and CTU were not significant.

Table 4. Driving performance factors under the ND and NC conditions.

Factors		Max.	Min.	Mean	SD	Abs. diff. (% diff.)	*SE*	*F*-statistics (P-value)
BRT	ND	1.150	0.501	0.727	0.149	0.156	0.027	5.965
(s)	NC	1.867	0.499	0.883	0.315	(21.5%)	0.059	(0.018)
DSF	ND	1.421	0.095	0.823	0.335	0.245	0.611	5.311
(km/h)	NC	2.073	0.243	1.068	0.471	(29.8%)	0.089	(0.025)
CDU	ND	8.978	0.756	2.967	1.567	0.207	0.286	0.398
(m)	NC	6.249	0.788	3.174	1.402	(7.0%)	0.264	(0.531)
CTU	ND	0.737	0.08	0.262	0.118	0.039	0.022	1.208
(s)	NC	0.727	0.03	0.301	0.15	(14.9%)	0.028	(0.276)

BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.
ND: no distraction; NC: normal conversation.

In summary, all factors were not significant (at the 1% level). However, brake reaction time was increased and driving speed stability was significantly worse under the normal conversation condition compared with the no distraction condition when driving, at the 5% level of significance.

4.2 Calculation levels versus no distraction

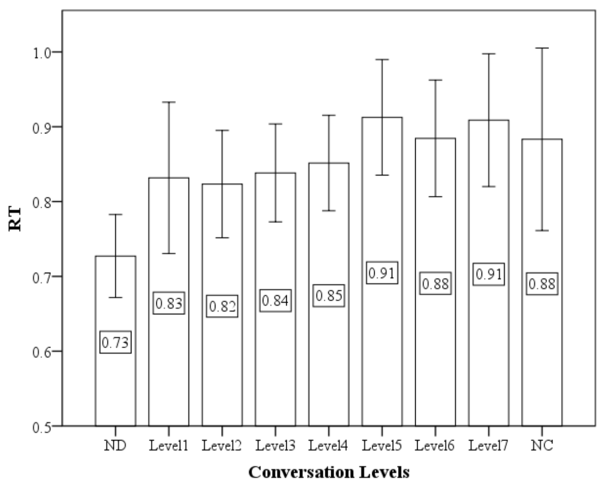
ANOVA was conducted to examine the differences in BRT, DSF, CDU, and CTU under the no distraction condition and each of the seven calculation task conditions. The ANOVA results of the differences in driving maneuver stability variables in the no distraction condition and the seven difficulty levels of calculation are shown in Table 5.

Table 5. Descriptive statistics of driving stability variables and ANOVA results.

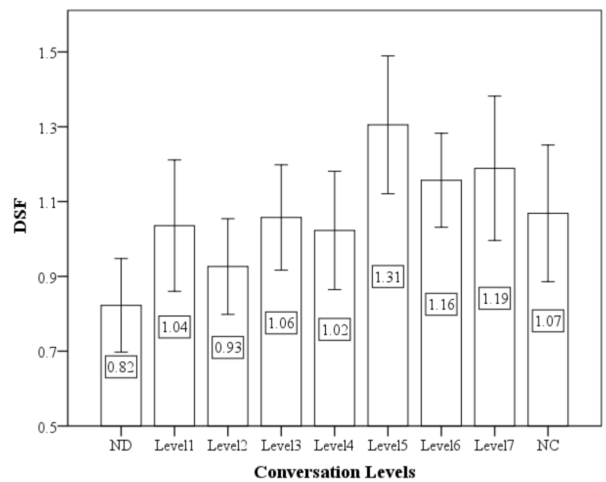
Distraction condition	BRT					DSF					CDU					CTU				
	Mean	SD	SE	Abs. diff. (% diff.)	*F*(P)	Mean	SD	SE	Abs. diff. (% diff.)	*F*(P)	Mean	SD	SE	Abs. diff. (% diff.)	*F*(P)	Mean	SD	SE	Abs. diff. (% diff.)	*F*(P)
No distraction	0.727	0.149	0.027	Base	--	0.823	0.335	0.061	Base	--	2.967	1.567	0.286	Base	--	0.262	0.118	0.022	Base	--
Level 1	0.832	0.271	0.049	0.105 (14.4 %)	5.431 (0.049)	1.036	0.471	0.086	0.213 (25.9 %)	4.087 (0.049)	3.725	1.568	0.286	0.758 (25.5 %)	3.887 (0.079)	0.33	0.141	0.026	0.068 (26.0 %)	4.029 (0.058)
Level 2	0.823	0.192	0.035	0.096 (13.2 %)	4.711 (0.07)	0.926	0.343	0.063	0.103 (12.5 %)	1.405 (0.337)	2.969	1.189	0.217	0.002 (0.1%)	0.014 (0.925)	0.272	0.086	0.016	0.010 (3.8%)	0.129 (0.787)
Level 3	0.838	0.175	0.032	0.111 (15.3 %)	7.015 (0.037)	1.058	0.377	0.069	0.235 (28.6 %)	6.510 (0.030)	3.488	2.178	0.398	0.521 (17.6 %)	1.312 (0.215)	0.319	0.172	0.031	0.057 (21.8 %)	2.234 (0.109)
Level 4	0.852	0.171	0.031	0.124 (17.1 %)	9.040 (0.019)	1.023	0.424	0.077	0.200 (24.3 %)	3.125 (0.065)	3.359	1.241	0.227	0.392 (13.2 %)	1.405 (0.339)	0.321	0.117	0.021	0.059 (22.5 %)	3.802 (0.095)
Level 5	0.913	0.207	0.038	0.183 (25.1 %)	15.898 (0.001)	1.305	0.494	0.901	0.482 (58.6 %)	19.650 (<0.001)	4.356	2.211	0.404	1.389 (46.8 %)	8.345 (0.002)	0.377	0.14	0.026	0.115 (43.9 %)	11.689 (0.001)
Level 6	0.884	0.209	0.038	0.157 (21.6 %)	11.304 (0.003)	1.157	0.336	0.061	0.334 (40.6 %)	14.905 (0.002)	3.903	1.41	0.257	0.936 (31.5 %)	6.432 (0.032)	0.348	0.132	0.024	0.086 (32.8 %)	7.089 (0.015)
Level 7	0.909	0.238	0.043	0.182 (25.0 %)	12.598 (0.001)	1.189	0.517	0.094	0.366 (44.5 %)	10.624 (0.001)	4.061	2.234	0.408	1.094 (36.9 %)	5.187 (0.013)	0.374	0.167	0.031	0.112 (42.7 %)	8.982 (0.002)

BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.

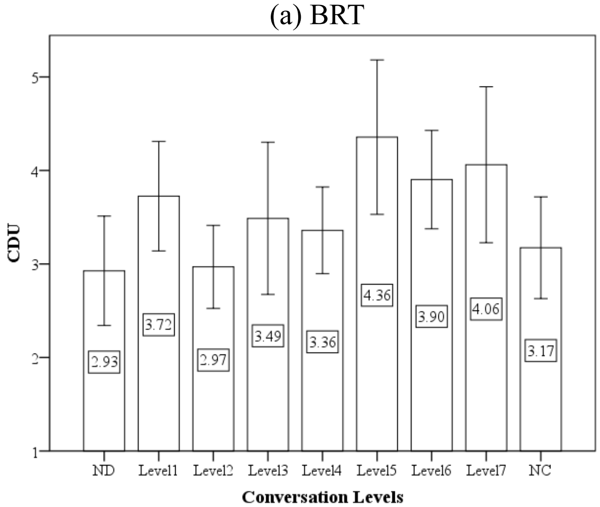
Significant differences in BRT were found between the no distraction condition and the calculation difficulties of Level 5, 6, 7 at the 1% significance level. For the Level 2, 3, 5 conditions, the differences were potential significant (at the 5% level). For DSF, the differences between no distraction and Levels 1, 3, and 4 were potential significant (at the 5% level), and those for Levels 5 to 7 at the 1% significance level. There was significant differences in CDU between the no distraction condition at just Levels 5 calculation conditions at the 1% level. And the Levels 6 and 7 conditions were potential significant (at the 5% level). Potential significant differences in CTU were found between the no distraction condition and the Level 6 condition (at the 5% significance level), and between the no distraction condition and the Levels 5 to 7 conditions at the 1% significance level. The average CTUs under the Levels 5 and 7 conditions were all higher than under the no distraction condition, with absolute differences ranging from 0.010 to 0.115 s and relative differences ranging from 3.8% to 43.9%. Figure 3 illustrates the average values of BRT, DSF, CDU, and CTU at different calculation levels comparing to which had no distraction.



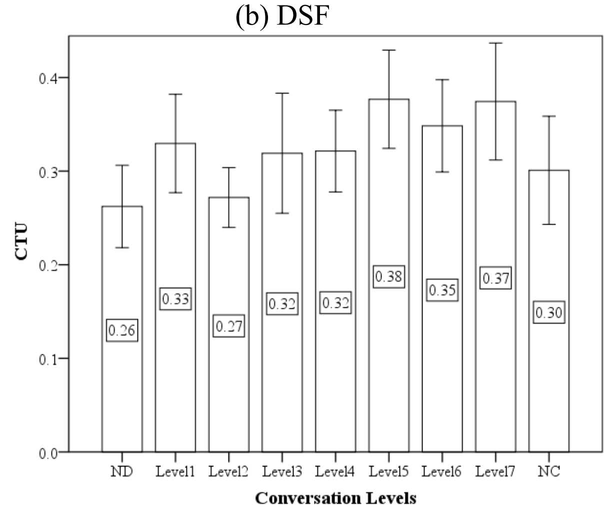
(a) BRT



(b) DSF



(c) CDU



(d) CTU

Figure 3. Mean BRT, DSF, CDU, and CTU under different calculation levels.

ND: no distraction.

The participants' driving maneuver stability variables of BRT, DSF, CDU, and CTU increased under all of the mathematical calculation task conditions and under the normal conversation condition compared with the no distraction condition. The measurements were not linearly correlated with the increase in the calculation difficulty levels; however, the re-divided seven calculation conditions of interval groups can be used to benchmark the conversation distraction levels.

4.3 Demographic

Standard one-way ANOVA *F*-statistics were used to evaluate the effects of demographic characteristics on driving performance factors. The results are presented in Table 6. To ensure sufficient sample sizes, the driving maneuver variables used to measure stability were combined under all of the conditions, including the ND, NC, and seven calculation conditions, rather than being separately analyzed.

Table 6. One-way ANOVA of driving performance factors and driver characteristics

F-statistics	Gender	Age	Years license held	Driving frequency	Driving duration	Cellular phone use habit
BRT	0.003	0.582	0.007	0.977	0.910	1.421
DSF	0.283	3.465	7.819**	0.206	0.880	0.391
CDU	0.232	0.506	9.274**	7.599**	7.469**	0.168
CTU	0.937	0.059	7.626**	2.284	4.326*	0.081

**means significant at 0.01 level
*means potential significant at 0.05 level
BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.

As shown in Table 6, the number of years that a person had held a full driving license was found significantly affect DSF, CDU, and CTU, which are measures of maneuver stability; all of the differences were significant at the 1% level. Driving frequency significantly affected CDU at the 1% significance level, and driving duration per week significantly affected CDU, which are again measures of maneuver stability, at the 1% significance level, respectively. However, none of the other demographic characteristics significantly affected the driving performance factors. The participants' gender, age, and cellular phone experience were not associated with significant differences in BRT, DSF, CDU, or CTU. BRT was not affected by any of the driver characteristics analyzed.

4.4 Carry calculation versus non-carry calculation

In this experiment design, the difficulty level of carry calculation is greater than non-carry calculation, and the calculation difficulty was hypothesized to increase from Level 1 to Level 7. However, BRT, DSF, CDU, and CTU were not found to consistently increase from Level 1 to Level 7. The effects of difficulty on BRT, DSF, CDU, and CTU under carry calculation were generally lower than under non-carry calculation, as shown in Figure 5, but the ANOVA results showed that these differences were not significant (BRT: p = 0.835; DSF: p = 0.334; CDU: p = 0.260; CTU: p =0.516). The carry calculation group and non-carry calculation group were then analyzed separately (see Figure 4). The carry calculation group consisted of Level 1, Level 3, and Level 5, and the non-carry calculations included Level 2, Level 4, Level 6, and Level 7. In the non-carry calculation group, the mean BRT and mean DSF displayed a trend of constant increase, but the mean CDU and CTU did not. In the carry calculation group, all four variables increased constantly from Level 2 to Level 7.

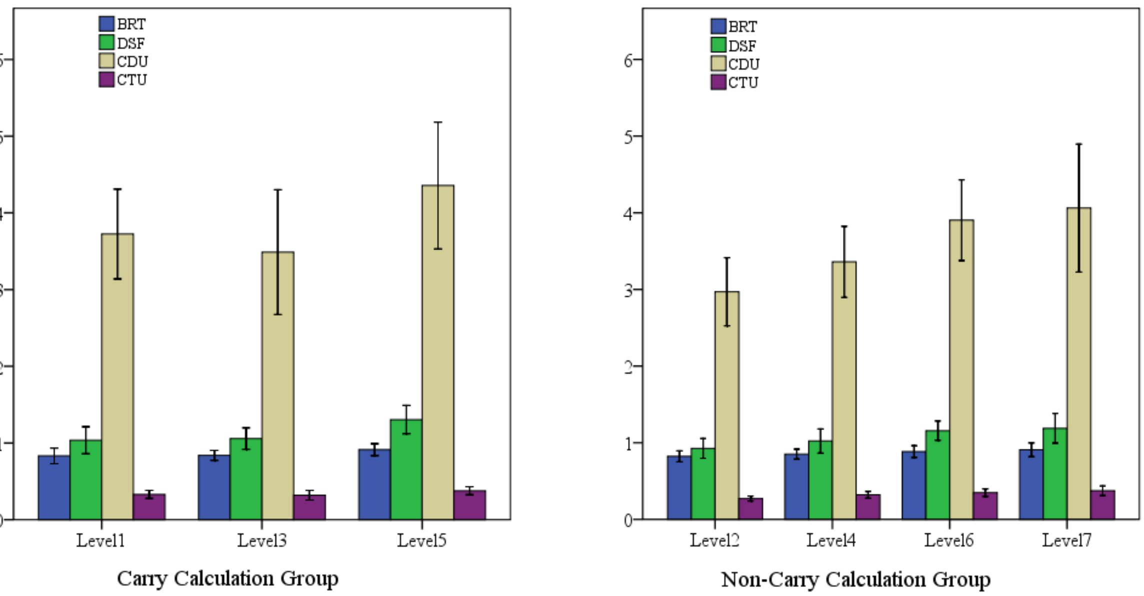


Figure 4: Mean BRT, DSF, CDU, and CTU under different calculation groups

4.5 Discriminant analysis of the calculation and normal conversation conditions

The Kruskal-Wallis and median tests were conducted in a weighted model to compare the effects of normal conversation with the effects of the calculation tasks on driving performance.

4.5.1 *Kruskal-Wallis test*

Table 7. Ranks.

Mean rank		Conversational calculation level							NC
		Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	
Driving performance factors	BRT	121.68	123.83	133.75	142.23	160.98	150.43	155.50	138.88
	DSF	124.37	110.73	134.73	128.00	174.53	156.43	151.40	136.75
	CDU	145.17	109.83	124.70	133.43	166.20	157.93	150.07	119.79
	CTU	136.30	106.33	130.57	141.03	167.03	151.50	159.97	120.36

BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.
NC: normal conversation.

The Kruskal-Wallis test was used to conduct ANOVA based on ranks. The absolute differences in ranks between the normal conversation condition and the mathematical calculation levels were calculated based on the weighting 3:1:1:1 (BRT: DSF: CDU: CTU), and the overall absolute rank difference by weight was calculated. The results are shown in Table 8.

Table 8. Absolute differences in ranks and overall weighted rank difference.

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7
BRT	17.19	15.04	5.13	3.36	22.11	11.56	16.63
DSF	12.38	26.02	2.02	8.75	37.78	19.68	14.65
CDU	25.38	9.95	4.91	13.65	46.41	38.15	30.28
CTU	15.94	14.02	10.21	20.68	46.68	31.14	39.61
Overall weighted rank difference	105.28	95.12	32.52	53.15	197.20	123.65	134.42

BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.

The smaller the weighted rank difference between two distributions, the higher the similarity between the normal conversation condition and the specified calculation task level. As shown in Table 8, the smallest absolute differences in ranks for DSF, CDU, and CTU occurred in the Level 3 mathematical calculation condition. The absolute differences of ranks for BRT were smallest in the Levels 3 and 4 conditions. The overall rank absolute difference was smallest for the Level 3 condition. The results illustrated that the distributions under the normal conversation condition were similar to the distributions under the Level 3 condition in terms of the individual driving performance factors and the overall rank weight.

4.5.2 *Median test*

The median test was used to examine the frequency ratio of BRT, DSF, CDU, and CTU under the seven calculation task conditions. The results are presented in Tables 9 and 10.

Table 9. Frequency of median test.

Frequency		Conversational calculation level							NC
		Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	
BRT	> Median	14	11	15	17	21	18	17	15
	≤ Median	16	19	15	13	9	12	13	13
DSF	> Median	13	10	15	13	21	19	18	13
	≤ Median	17	20	15	17	9	11	12	15
CDU	> Median	18	11	12	17	20	21	16	11
	≤ Median	12	19	18	13	10	9	14	17
CTU	> Median	15	10	17	17	21	17	18	11
	≤ Median	15	20	13	13	9	13	12	17

BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.
NC: normal conversation.

Table 10. Rate of median test.

Rate		Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	NC
BRT	> Median	46.7%	36.7%	50.0%	56.7%	70.0%	60.0%	56.7%	53.6%
	≤ Median	53.3%	63.3%	50.0%	43.3%	30.0%	40.0%	43.3%	46.4%
DSF	> Median	43.3%	33.3%	50.0%	43.3%	70.0%	63.3%	60.0%	46.4%
	≤ Median	56.7%	66.7%	50.0%	56.7%	30.0%	36.7%	40.0%	53.6%
CDU	> Median	60.0%	36.7%	40.0%	56.7%	66.7%	70.0%	53.3%	39.3%
	≤ Median	40.0%	63.3%	60.0%	43.3%	33.3%	30.0%	46.7%	60.7%
CTU	> Median	50.0%	33.3%	56.7%	56.7%	70.0%	56.7%	60.0%	39.3%
	≤ Median	50.0%	66.7%	43.3%	43.3%	30.0%	43.3%	40.0%	60.7%

BRT: brake reaction time; DSF: driving speed fluctuation; CDU: car-following distance undulation; CTU: car-following time headway undulation.
NC: normal conversation.

The absolute differences in the median rates under the normal conversation and mathematical calculation level conditions were calculated based on the 3:1:1:1 (BRT: DSF: CDU: CTU) weighting. Table 11 shows the overall absolute median rate differences by weight.

Table 11. Absolute differences of median rates and overall weighted absolute difference.

Abs. diff	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7
RT	13.8%	33.8%	7.1%	6.2%	32.9%	12.9%	6.2%
DSF	6.2%	26.2%	7.1%	6.2%	47.1%	33.8%	27.1%
CDU	41.4%	5.2%	1.4%	34.8%	54.8%	61.4%	28.1%
CTU	21.4%	11.9%	34.8%	34.8%	61.4%	34.8%	41.4%
Overall weighted median rate difference	110.5%	144.8%	64.8%	94.3%	261.9%	168.6%	115.2%

The smaller the median rate difference, the greater the similarity between the normal conversation condition and the specified mathematical calculation task condition. The absolute differences in the median rates of DSF and CDU were smallest under the Level 3 condition. The absolute difference in the median rate of CTU was smallest in the Level 2 condition. The absolute differences in the median rates of BRT were smallest under the Level 3, 4, and 7 conditions. The overall rank absolute median rate was smaller under the Level 3 condition than under the other conditions. This illustrates that the median rate of distributions under the normal conversation condition was most similar to that under the Level 3 condition in terms of the overall rank weight.

The Kruskal-Wallis and median test analyses revealed that the distraction caused by normal conversation was comparable in degree to the distraction caused by calculation tasks. After analyzing the individual driving factors and the overall weight measurements, we found that the distraction caused by normal conversation was most similar to that caused by performing Level 3 calculations.

5 DISCUSSION

5.1 Impairment effects of normal conversation while driving

This study demonstrated that driving performance measures, including BRT and DSF, were significantly impaired by normal conversation on hands-free cellular phones. These results were consistent with findings that engaging in hands-free cellular phone conversations while driving generally impairs driving performance (Charlton, 2009, Caird et al., 2008, Rosenbloom, 2006).

Although in our study, hands-free cellular phone conversation has no significant effect on BRT, other numerous studies have concluded that hands-free cellular phone conversations have significant effects on BRT; these include simulation studies (Charlton, 2009), an observation study conducted in a real driving environment, and comprehensive studies (Caird et al., 2008, Horrey and Wickens, 2006, Al-Darrab et al., 2009, Lamble et al., 1999). This study found that engaging in hands-free cellular phone conversations while driving increased drivers' BRT by 21.5%. This is consistent with studies that found an increase of 15% to 40% in BRT when participants used a cell phone while driving. DSF has also been used to test the effect of hands-free cellular conversations on driving performance (Rosenbloom, 2006, Charlton, 2009, Treffner and Barrett, 2004, Saifuzzaman et al., 2015, Lamble et al., 1999). Some studies have found that CDU is significantly impaired by hands-free cellular phone conversations (Rosenbloom, 2006, Charlton, 2009, Saifuzzaman et al., 2015, Lamble et al., 1999, Atchley and Dressel, 2004). In this study, although CDU was not significantly impaired by engaging in calculation tasks, it was still included in the analysis. With a larger sample size, the effects of distraction on these driving performance indicators could become significant.

5.2 Measurements of distraction due to normal conversation

In this study, seven-level mathematical calculations were used to induce various levels of distraction in mental workload and were compared with the distraction caused by normal conversation. K-S and median tests both verified that the distraction caused by normal conversation was similar to that occurring under the Level 3 calculation task condition. The level of impairment was measured using BRT, DSF, CDU, and CTU, and a weight-based model that combined the four factors.

Some studies have used verbal recall or recognition tasks (e.g., listening to sentences, remembering elements of the sentences, and then repeating the words or making some sort of decision about the words) as distraction tasks (Haigney et al., 2000, Mazzae et al., 2004, Strayer and Johnston, 2001). Other studies have used naturalistic phone conversations to measure the effects of cellular phone usage (Rakauskas et al., 2004, Shinar et al., 2005). However, these measures are neither clear nor comparable. To standardize the distraction levels, we conducted the standard and replicable methods as some studies - solve mathematical problems (Brookhuis et al., 1991, McKnight and McKnight, 1993, Patten et al., 2004).

Studies using mathematical problems as standardized driving distractions have examined the distraction effects of simple and complex phone conversations on driving performance. In a simulated study, McKnight and McKnight (1993) used a sample of 150 participants to examine five distraction conditions while driving: dialing a cellular phone, holding a simple phone conversation (e.g., discussing what they did for a living), holding a complex phone conversation (e.g., solving a math problem), tuning a radio, and no distraction. All three phone conditions led to an increase in failures to respond to traffic situations, and the complex conversation led to the most seriously impaired driving performance. That study represented the first use of mathematical calculations in distraction analysis studies (McKnight and McKnight, 1993). Patten et al. (2004) investigated the effects of different distraction tasks on driving performance: (1) an ND condition, (2) a simple conversation requiring drivers to repeat back single digits spoken by the experimenter, and (3) solving arithmetic problems. The study found that solving arithmetic problems had a more negative effect on drivers' reactions to a peripheral detection task than having a simple conversation (Patten et al., 2004). Harbluk et al. (2002) investigated the effect of cognitive distraction on driving in an on-road experiment; they specified the calculation difficulty levels of easy addition problems (e.g., 6 + 9) and complex addition tasks (e.g., 47 + 38). As the complexity of the addition tasks increased, the drivers' perceptions of workload, distraction level, and perceptions of their driving as being less safe all increased (Harbluk et al., 2002); their vehicle control and their inspection glances to traffic lights at intersections were also impaired (Harbluk et al., 2007).

Our seven-level calculation scheme offered a refined classification of mathematical calculation levels and provided a method for inducing various levels of task distraction in a standard system. As the comparison of various levels of distraction is a common study method, the seven-level calculation model could serve as a standardized measure of driving distraction tasks.

5.3 Effect of driving experience on driving performance

This study found that driving experience affected driving performance. Baker-Grøndahl and Sagberg examined how the number of years a participant had held a driving license affected driving performance. They found that having a driving license for a long time may falsely influence drivers' self-evaluations; they may be too confident about their driving performance (Backer-Grøndahl and Sagberg, 2011). Because, in Hong Kong, many more people hold a driving license than own vehicles or drive frequently, having a driving license may not represent real driving experience and may even lead to incorrect evaluations.

However, increased frequency and duration of driving improves driver performance. The participants who drove regularly had better driving maneuver capability. In this study, the participants who drove more than once a week performed better in terms of car following stability. In terms of driving frequency, participants with 3 hours and longer driving duration had better car stability than those who drove once a week and those who drove 3 hours or less a week. Thus, training and regular practice lead to safer vehicle maneuvers.

5.4 Benchmark for distraction assessment by carry calculation versus non-carry calculation

At the experimental design stage, seven levels of mathematical calculations were designed, such that the brake reaction time should increase linearly according to increasing difficulty level, and the impairment of drivers' maneuver stability should gradually increase with increasing difficulty level. However, according to the analysis of the results, the difficulty of the mathematical calculations did not linearly increase from Level 1 to Level 7; rather, the linear increase occurred at two-level intervals, so that the distraction benchmark could be grouped to include either Level 2, Level 4, Level 6, and Level 7, or Level 1, Level 3, and Level 5. Hence, these mathematical calculations were regrouped and standardized.

The regrouped set of calculation difficulty levels comprising Level 2, Level 4, Level 6, and Level 7 involves the calculation of a decimal carry-unit digit, for which the result requires at least one digit addition carry. The other regrouped benchmark, consisting of Level 1, Level 3, and Level 5, involves the calculation of no-addition carry, which can be considered as two non-addition carries conducted in the unit-digit and ten-digit, respectively. Both of these regrouped mathematical calculation benchmarks are applicable. Hence, we can still conclude that the impairment of the drivers' maneuver stability induced by distraction can be assessed by the seven-level calculation benchmark and that the impairment increases along with increasing difficulty level.

5.5 Study limitations

Studies using the driving simulator approach have covered various topics, however, driving simulation has some unavoidable disadvantages that may possibly confound statistics analysis (Lerman et al., 1993), limit the effectiveness of training (Wu et al., 2016), and lead to overconfident driving intentions of novice drivers in performing unsafe driving (Rosenbloom and Eldror, 2014), breed familiarity and inattention by repeated simulating in simplified real route conditions (Yanko and Spalek, 2013), driving performance with and represent an unreal driving situation (Yan et al., 2016). Our study could also not avoid these defects as mentioned above. However, the trend research of driving behavior is still feasible. When having a hands-free conversation while driving, self-regulation or behavioral adaptation in the simulated driving experiment can lead to the similar engaging behaviors in the daily life. Therefore, the task of this study requires drivers to keep a specified speed. If the difference between driver's speed and specified speed exceed 30%, the experiment failed and data was unusable. Self-regulation is feasible in this study. In addition, the participant sample could be extended to larger scales.

Yet using a driving simulator is still the most feasible instrument for comprehensive studies using sophisticated experiments designed to isolate specific relationships between different factors influencing driving behavior, particularly when road safety is uncertain in some conditions.

6 CONCLUSIONS

This driving simulator study of driver's distraction by hands-free cellular phone conversation examined the effects of normal conversation on drivers, using seven-level difficulty mathematical calculations as a benchmark to assess distraction-induced impairment. The distraction effects of normal conversation were compared with the standardized levels of distraction caused by various levels of calculation. Four driving performance indicators of maneuver stability, including BRT, DSF, CDU, and CTU were used to measure and compare the drivers' vehicle maneuver capability under three conditions: (1) no distraction, (2) normal conversation, and (3) seven-level mathematical calculation. The distraction caused by a normal conversation on a cellular phone was significant.

Thirty participants, half of whom had previous experience using cellular phones while driving, took part in the simulated driving experiment. The results were as follows.

(1) Normal conversation impaired driving performance. The standardized seven-level calculation tasks could be regrouped by the presence or absence of digit carry, and the group comprising Levels 2, 4, 6, and 7, involving the calculation of a decimal carry-unit digit, can be used to measure the effects of normal conversation on driving performance.
(2) The distraction caused by normal conversation was equivalent to the distraction caused by mathematical (addition) calculation of Level 3. This level involved the addition of a single- and a double-digit number, such as 31 + 6 or 84 + 3, but did not require the calculation of a decimal carry-unit digit or a carry-10 digit.
(3) Driving experience has a positive effect on driving stability; however, the number of years a participant has held a driving license has a negative effect on maneuver stability for the young occasional drivers.

This study has implications for road safety enforcement applications and educational strategies. In terms of enforcement, it supports penalties to deter hands-free cellular phone conversations, because they significantly impair many aspects of driving performance. In terms of education, drivers should be given a clear sense of the distraction caused by normal conversation. It is important to highlight that even hands-free cellular phone conversations impair driving performance. The seven-level mathematical calculation tasks can be used in studies to assess and compare driving distraction impairment. Future studies could explore the impairment associated with various distracted driving conditions, such as emotional engagement and conversations with in-car passengers. Other types of mathematical calculation benchmarks could also be developed for studies of distracted driving performance.

ACKNOWLEDGEMENTS

The research described here was supported by a Research Postgraduate Studentship, and grants from the University Research Committee of the University of Hong Kong (201511159015), the Joint Research Scheme of National Natural Science Foundation of China / Research Grants Council of Hong Kong (Project No. 71561167001 & N_HKU707/15), the National Natural Science Foundation of China (71701023), the Outstanding Youth Foundation of Hunan Education Department (17B015), and Open Fund of Key Laboratory of Special Environment Road Engineering of Hunan Province (Changsha University of Science & Technology) (KFJ160503). We thank Dan LU and Liqiao ZHANG for their help with the data collection.

REFERENCES

1. AL-DARRAB, I. A., KHAN, Z. A. & ISHRAT, S. I. 2009. An experimental study on the effect of mobile phone conversation on drivers' reaction time in braking response. *Journal of Safety Research*, 40, 185-189.

2. ALM, H. & NILSSON, L. 1994. Changes in driver behaviour as a function of handsfree mobile phones - A simulator study. *Accident Analysis & Prevention*, 26, 441-451.
3. ALM, H. & NILSSON, L. 1995. The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis & Prevention*, 27, 707-715.
4. ATCHLEY, P. & DRESSEL, J. 2004. Conversation limits the functional field of view. *Human Factors*, 46, 664-673.
5. BACKER-GRØNDAHL, A. & SAGBERG, F. 2011. Driving and telephoning: Relative accident risk when using hand-held and hands-free mobile phones. *Safety Science*, 49, 324-330.
6. BEEDE, K. E. & KASS, S. J. 2006. Engrossed in conversation: The impact of cell phones on simulated driving performance. *Accident Analysis & Prevention*, 38, 415-421.
7. BROOKHUIS, K. A., DE VRIES, G. & DE WAARD, D. 1991. The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention*, 23, 309-316.
8. CAIRD, J. K., WILLNESS, C. R., STEEL, P. & SCIALFA, C. 2008. A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention*, 40, 1282-1293.
9. CAIRD, J. K., JOHNSTON, K. A., WILLNESS, C. R., ASBRIDGE, M. & STEEL, P. 2014. A meta-analysis of the effects of texting on driving. *Accident Analysis & Prevention*, 71, 311-318.
10. CHARLTON, S. G. 2009. Driving while conversing: cell phones that distract and passengers who react. *Accident Analysis & Prevention*, 41, 160-173.
11. COOPER, P. J., ZHENG, Y., RICHARD, C., VAVRIK, J., HEINRICHS, B. & SIEGMUND, G. 2003. The impact of hands-free message reception/response on driving task performance. *Accident Analysis & Prevention*, 35, 23-35.
12. CUNNINGHAM, M. L., REGAN, M. A., & IMBERGER, K. 2017. Understanding driver distraction associated with specific behavioural interactions with in-vehicle and portable technologies. *Journal of the Australasian College of Road Safety*, 28(1), 27.
13. DREWS, F. A., PASUPATHI, M. & STRAYER, D. L. 2008. Passenger and cell phone conversations in simulated driving. *Journal of Experimental Psychology: Applied*, 14, 392-400.
14. FITCH, G. M., BARTHOLOMEW, P. R., HANOWSKI, R. J. & PEREZ, M. A. 2015. Drivers' visual behavior when using handheld and hands-free cell phones. *Journal of Safety Research*, 54, 105.e29-108.
15. FITCH, G. A., SOCCOLICH, S. A., GUO, F., MCCLAFFERTY, J., FANG, Y., OLSON, R. L., PEREZ, M. A., HANOWSKI, R. J., HANKEY, J. M., & DINGUS, T. A. (2013, April). The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk. (Report No. DOT HS 811 757). Washington, DC: National Highway Traffic Safety Administration.
16. HANCOCK, P. A., LESCH, M. & SIMMONS, L. 2003. The distraction effects of phone use during a crucial driving maneuver. Accident Analysis & Prevention, 35, 501-514.
17. HAIGNEY, D. E., TAYLOR, R. G. & WESTERMAN, S. J. 2000. Concurrent mobile (cellular) phone use and driving performance: task demand characteristics and compensatory processes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 3, 113-121.
18. HARBLUK, J. L., NOY, Y. I. & EIZENMAN, M. 2002. The impact of cognitive distraction on driver visual behaviour and vehicle control (TP No. 13889 E). Ottawa, Canada: Transport Canada.

19. HARBLUK, J. L., NOY, Y. I., TRBOVICH, P. L. & EIZENMAN, M. 2007. An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accident Analysis & Prevention*, 39, 372-379.
20. HORREY, W. J. & WICKENS, C. D. 2006. Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors*, 48, 196-205.
21. ISHIGAMI, Y. & KLEIN, R. M. 2009. Is a hands-free phone safer than a handheld phone? *Journal of Safety Research*, 40, 157-164.
22. LAMBLE, D., KAURANEN, T., LAAKSO, M. & SUMMALA, H. 1999. Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving. *Accident Analysis & Prevention*, 31, 617-623.
23. LERMAN, Y., SADOVSKY, G., GOLDBERG, E., KEDEM, R., PERITZ, E. & PINES, A. 1993. Correlates of military tank simulator sickness. *Aviation, Space, and Environmental Medicine*, 64**,** 619-22.
24. Li, X., Yan, X., Wu, J., Radwan, E., & Zhang, Y. (2016). A rear-end collision risk assessment model based on drivers' collision avoidance process under influences of cell phone use and gender—A driving simulator based study. *Accident Analysis & Prevention*, 97, 1-18.
25. Lipovaca, K., Đerićba, M., Tešićac, M., Andrićb, Z., Marićd, B. 2017. Mobile phone use while driving-literary review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 47, 132-142.
26. LIU, B. S. & LEE, Y.H. 2005. Effects of car-phone use and aggressive disposition during critical driving maneuvers. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8, 369-382.
27. LIU, B. S. & LEE, Y. H. 2006. In-vehicle workload assessment: Effects of traffic situations and cellular telephone use. *Journal of Safety Research*, 37, 99-105.
28. MACIEJ, J., NITSCH, M. & VOLLRATH, M. 2011. Conversing while driving: The importance of visual information for conversation modulation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14, 512-524.
29. MAZZAE, E. N., RANNEY, T. A., WATSON, G. S. & WIGHTMANN, J. A. 2004. Hand-held or hands-free? The effects of wireless phone interface type on phone task performance and driver preference. *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society*, 2218-2222.
30. MCKNIGHT, A. J. & MCKNIGHT, A. S. 1993. The effect of cellular phone use upon driver attention. *Accident Analysis & Prevention*, 25, 259-265.
31. NHTSA. 2016. *Visual-manual NHTSA driver distraction guidelines for portable and aftermarket devices*. Washington, DC: National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT).
32. OVIEDO-TRESPALACIOS, O., HAQUE, M. M., KING, M., & WASHINGTON, S. (2016). Understanding the impacts of mobile phone distraction on driving performance: a systematic review. *Transportation research part C: emerging technologies*, 72, 360-380.
33. PATTEN, C. J. D., KIRCHER, A., ÖSTLUND, J. & NILSSON, L. 2004. Using mobile telephones: cognitive workload and attention resource allocation. *Accident Analysis & Prevention*, 36, 341-350.
34. RAKAUSKAS, M. E., GUGERTY, L. J. & WARD, N. J. 2004. Effects of naturalistic cell phone conversations on driving performance. *Journal of Safety Research*, 35, 453-464.

35. ROAD SAFETY COUNCIL 2003. *Hong Kong Road Safety Review*. Road Safety Research Committee. December 2003. Hong Kong SAR.
36. ROSENBLOOM, T. 2006. Driving performance while using cell phones: an observational study. *Journal of Safety Research,* 37, 207-212.
37. ROSENBLOOM, T. & ELDROR, E. 2014. Effectiveness evaluation of simulative workshops for newly licensed drivers. *Accident Analysis & Prevention*, 63, 30-36.
38. SAIFUZZAMAN, M., HAQUE, M. M., ZHENG, Z. & WASHINGTON, S. 2015. Impact of mobile phone use on car-following behaviour of young drivers. *Accident Analysis & Prevention*, 82, 10-19.
39. SHINAR, D., TRACTINSKY, N. & COMPTON, R. 2005. Effects of practice, age, and task demands, on interference from a phone task while driving. *Accident Analysis & Prevention*, 37, 315-326.
40. SIMMONS, S. M., HICKS, A. & CAIRD, J. K. 2016. Safety-critical event risk associated with cell phone tasks as measured in naturalistic driving studies: A systematic review and meta-analysis. *Accident Analysis & Prevention*, 87, 161-169.
41. STRAYER, D. L. & DREWS, F. A. 2004. Profiles in driver distraction: effects of cell phone conversations on younger and older drivers. *Human Factors*, 46, 640-649.
42. STRAYER, D. L., DREWS, F. A. & JOHNSTON, W. A. 2003. Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, 9, 23-32.
43. STRAYER, D. L. & JOHNSTON, W. A. 2001. Driven to distraction: dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science (Wiley-Blackwell)*, 12, 462.
44. SUNYAEV, S. R., EISENHABER, F., ARGOS, P., KUZNETSOV, E. N. & TUMANYAN, V. G. 1998. Are knowledge-based potentials derived from protein structure sets discriminative with respect to amino acid types? *Proteins: Structure, Function, and Bioinformatics*, 31, 225-246.
45. THAPA, R., CODJOE, J., ISHAK, S., MCCARTER, K. S. 2015. Post and During Event Effect of Cell Phone Talking and Texting on Driving Performance -A Driving Simulator Study. *Traffic Injury Prevention*, 5, 1-7.
46. TÖRNROS, J. E. B. & BOLLING, A. K. 2005. Mobile phone use—effects of handheld and handsfree phones on driving performance. *Accident Analysis & Prevention*, 37, 902-909.
47. TREFFNER, P. J. & BARRETT, R. 2004. Hands-free mobile phone speech while driving degrades coordination and control. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7, 229-246.
48. TRIVEDI, N., HAYNIE, D., BIBLE, J., LIU, D., SIMONS-MORTON, B. 2017. Cell Phone Use While Driving: Prospective Association with Emerging Adult Use. *Accident Analysis and Prevention*, 106, 450-455.
49. WHO. 2011. Mobile phone use: a growing problem of driver distraction. (http://www.who.int/violence_injury_prevention/publications/road_traffic/distracted_driving/en/)
50. Wu, J., Yan, X., & Radwan, E. (2016). Discrepancy analysis of driving performance of taxi drivers and non-professional drivers for red-light running violation and crash avoidance at intersections. Accident Analysis & Prevention, 91, 1-9.
51. YAN, W., WONG, S. C., LI, Y. C., SZE, N. N. & YAN, X. 2015. Young driver distraction by text messaging: a comparison of the effects of reading and typing text messages in

Chinese versus English. *Transportation Research Part F: Traffic Psychology and Behaviour*, 31, 87-98.
52. Yan, X., Zhang, X., Zhang, Y., Li, X., & Yang, Z. (2016). Changes in Drivers' Visual Performance during the Collision Avoidance Process as a Function of Different Field of Views at Intersections. PLoS one, 11(10), e0164101.
53. YANKO, M. R. & SPALEK, T. M. 2013. Route familiarity breeds inattention: A driving simulator study. *Accident Analysis & Prevention*, 57**,** 80-86.

APPENDIX

1. SAMPLES OF QUESTIONS FOR NORMAL CONVERSATION SESSION

How many family members do you have?
Where is your hometown?
How old are you?
What is your major?
What is your favorite food?
What is your favorite drink?
What is your favorite snack?
What is your favorite fruit?
What is your favorite pet?
Where have you been for travelling?

2. EXPLANATORY VARIABLES EQUATIONS

The calculation equations of participants' driving performances indicators are as follows:

(b) Driving speed fluctuation (DSF)

DSF measures the fluctuation in a test vehicle's speed and represents a driver's ability to follow a leading vehicle steadily. It is evaluated as the SD of the speed at different sampling times along the drive and can be expressed as follows:

$$\sigma_v = \sqrt{\frac{\sum_{i=1}^{N}\left(v_i - \overline{v}\right)^2}{N}},$$

where σ_v is the DSF, v_i is the speed of the test vehicle at sampling time i, and $\overline{v} = \frac{1}{N}\sum_{i=1}^{N} v_i$ is the average speed of the test vehicle during the sampling period.

(c) Car-following distance undulation (CDU)

CDU measures the fluctuating distance between a leading vehicle and a test vehicle and represents the test vehicle driver's ability to follow the leading vehicle, driving preference, and response to the leading vehicle's performance.

It is calculated as the SD of the car-following distance between the two vehicles at different sampling times along the drive and can be expressed as follows:

$$\sigma_d = \sqrt{\frac{\sum_{i=1}^{N}\left(d_i - \overline{d}\right)^2}{N}},$$

where σ_d is the CDU, d_i is the speed of the test vehicle at sampling time i, and $\overline{d} = \frac{1}{N}\sum_{i=1}^{N} d_i$ is the average car-following space between the test and leading vehicles during the sampling period.

(d) Car-following time headway undulation (CTU)

CTU measures the fluctuation in time headway between a leading car and a test car and represents the test car driver's ability to follow the leading car, based on his or her driving preference and response to the leading car's time headway measurement.

It is calculated as the SD of the car-following time headway between the leading vehicle and the test car at different sampling times along the drive and can be expressed as follows:

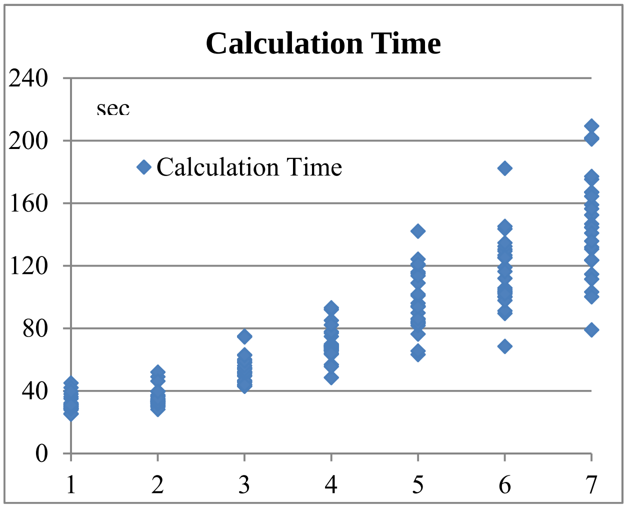
$$\sigma_t = \sqrt{\frac{\sum_{i=1}^{N}\left(t_i - \bar{t}\right)^2}{N}},$$

where σ_t is the CTU, t_i is the distance of the time headway of the test car at sampling time i, which is calculated from instantaneous brake reaction time data as $t_i = \frac{d_i}{v_i}$, and $\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_i$ is the average car-following time headway between the test car and the leading car during the sampling period.

3. VALIDATION OF MATHEMATICAL CALCULATION DIFFICULTY LEVELS

To test the validity and rationality of the difficulty levels and to evaluate each participant's mental workload for the mathematical calculation task, the average calculation time and accuracy of each level were collected and processed. Twenty addition calculations at each of the seven difficulty levels were posed to each participant via cellular phone. The hypothesis was that the required calculation time would increase linearly as the calculation level increased, that is, the number of calculation mistakes would comply with linear regressions based on the calculation difficulty levels.

In evaluating the mental workloads, the difficulty of the calculations was assessed based on the calculation time for each mathematical question and the overall percentage of accurate answers to the calculations. Figures 2 and 3 show plots of the calculation time and the number of mistakes, which both increased linearly with the seven levels of difficulty.



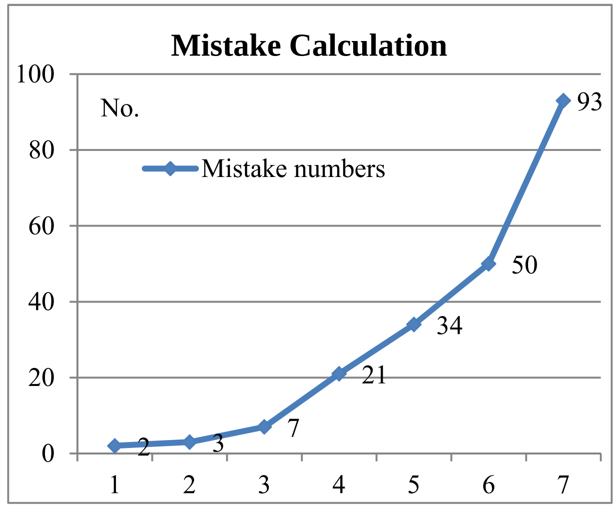


Figure 2 Calculation time by difficulty level Figure 3 Total number of mistakes by difficulty level

As shown in Figure 2, the calculation time increased as the level of difficulty increased. T represents the time taken by an individual to complete 20 mathematical calculations, and L_i represents the calculation level, with $i = 1, 2, 3, \cdots 6, \text{and } 7$ representing difficulty levels 1 to 7. β_0 and β_1 are the coefficients. The difficulty level model for calculation time is as follows:

$$T = \beta_0 + \beta_1 * L_i$$

Using an adjusted R^2 of 0.813, $\beta_0 = 0.044$, and $\beta_1 = 0.952$, the calculation time was positively correlated with the level of difficulty, as described by the following equation:

$$T = 0.044 + 0.952 * L_i$$

$$p < 0.01.$$

Thus, a unit increase in the level of difficulty resulted in a 0.952-s increase in calculation time. This behavior of the mathematical calculation time confirmed the reliability of the distraction level system in our model.

As shown in Figure 3, the total number of mistakes increased as the level of difficulty increased. M represents the overall mistakes by all participants in completing 20 mathematical calculations, and L_i represents the calculation level, with $i = 1, 2, 3, \cdots 6, \text{and } 7$ representing difficulty levels 1 to 7. β_0 and β_1 are the coefficients. The difficulty level model for accuracy is as follows:

$$M = \beta_0 + \beta_1 * L_i$$

Using an adjusted R^2 of 0.822, $\beta_0 = -26.286$, and $\beta_1 = 14.071$, the number of calculation mistakes was positively correlated with the level of difficulty, as described by the following equation:

$$T = \text{-}26.286 + 14.071 * L_i$$
$$p < 0.01.$$

Thus, a unit increase in the level of difficulty resulted in 14.071 additional mistakes. This behavior of the calculation accuracy again confirmed the reliability of the distraction level system in our model.

In summary, the calculation time and number of mistakes both increased linearly as the level of difficulty increased.