Risky Driving Behaviours Among Canadian Adults – A Cluster Analysis

by

Stephanie Hughes

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
February 2016

Table Of Contents

List Of Tables

List Of Figures

Abstract

Background: Motor vehicle collisions (MVCs) are a major cause of premature death and injury and are a major population health concern. Examination of MVCs and the role of human factors in traffic safety highlight the importance of clarifying the factors associated with risky driving behaviours (RDBs) and determining whether RDBs coexist within the same drivers so as to more effectively identify high-risk drivers. Research thus far has explored RDBs independently of one another, included a limited number of driver characteristics, investigated sub-populations not generalizable to the larger driving population, assessed RDBs in relation to MVC risk, and have varied in terminology, definition, and measurement of RDBs, thus limiting the interpretation of much of the current literature.

Objective: The objective was to explore the key factors associated with RDBs and to employ cluster analysis to determine whether distinct patterns of RDBs exist and, if so, identify the characteristics that differentiate subgroups of drivers. Associations between mental health factors and patterns of RDBs were also explored.

Methods: This study used secondary data from Canadian drivers 16 years and older who completed the Driving and Safety optional module of the 2011 Canadian Community Health Survey (CCHS). Descriptive analyses were used to explore whether associations existed between a large number of sociodemographic, health, and risk-taking variables and driver engagement in six RDBs. Ward's cluster analysis and *k*-means cluster analysis were employed to determine whether distinct patterns of driving behaviours existed within the driving population, and logistic regression models examined the characteristics that differentiated subgroups of drivers.

Results: Cluster analysis of RDBs revealed five heterogeneous clusters of drivers. Two subgroups of drivers, the Poly-risk Drivers (20.6%) and Egocentric Drivers (11.7%), demonstrated two very risky patterns of driving behaviours, while the subgroup of Average Drivers (30.0%), demonstrated a third moderately risky pattern of driving behaviours. Seat belt non-compliance was restricted to a small subset of drivers, the Beltless Drivers (4.4%). The Cautious Drivers (33.3%) refrained from all forms of RDBs. Descriptive characteristics uniquely profiled each subgroup of drivers and validated the five-cluster solution. Mental health factors were significantly associated with two patterns of very risky driving. Diagnosis of a mood disorder, higher levels of stress, and negative mental health were associated with a pattern of excessive speeding, severely aggressive driving, and fatigued driving (the Poly-risk Drivers); and 2. Higher levels of stress were associated with cell phone-distracted driving, driving under the influence of alcohol, and moderate speeding and aggressive driving (the Egocentric Drivers).

Conclusion: RDBs coexisted among diverse subgroups of drivers. External factors differentiated and profiled subgroups of drivers who engaged in multiple RDBs. Lifestyle, physical and mental health, and sociodemographic factors played a role in drivers who engaged in RDBs. The inclusion of such factors in future research may have important implications for understanding traffic safety.

List Of Abbreviations And Symbols Used

α	Alpha
ANOVA	Analysis of variance
BAC	Blood alcohol concentration
BAI	Beck Anxiety Inventory
β	Beta
CCHS	Canadian Community Health Survey (2011)
CCMTA	Canadian Council of Motor Transport Administration
CI	Confidence interval
DBQ	Driver Behaviour Questionnaire
$dist^2$	Distance squared
DNR	Did not report
DK/Ref	Don’t know/refuse
DSI	Driver Stress Inventory
DUIA	Driving under the influence of alcohol
ϵ	Element of
Freq	Frequency
F	F-test
γ	Gamma
GDL	Graduated Licensing Program
GHQ-12	General Health Questionnaire
M	Mean
MVC	Motor vehicle collision
N	Population
n	Sample population
N/A	Not applicable
NORP 2010	National Occupant Restraint Program
OECD	Organization for Economic Cooperation
OR	Odds ratio
p	p-value (calculated probability)
RDBs	Risky driving behaviours
RSS 2015	Road Safety Strategy 2015
RWDD	Riding with a drinking driver
SD	Standard deviation
SE	Standard error
SSE	Sum of squared error
$\sum$	Summation
STEPs	Periodic Selective Traffic Enforcement Program
STRID	Strategy to Reduced Impaired Driving
TSC	Traffic safety culture
Z	Z-score

Acknowledgements

I would like to express my sincere gratitude to my thesis advisor, Prof. Mark Asbridge, for his guidance, patience, and support. I would like to also thank my thesis committee, Dr. Don Langille and Dr. Kathleen MacPherson, for their excellent feedback and encouragement.

Heather Hobson, at the Atlantic Research Data Centre, was a wonderful contact and source of knowledge in my hours spent at the centre. Thank you for your support.

Chapter 1: Introduction

1.0 Motor Vehicle Collisions And Road Safety

Motor vehicle collisions (MVCs) result in over 1.2 million deaths and 20 to 50 million injuries worldwide each year (World Health Organization, 2009). Although global MVC-related morbidity and mortality rates have steadily declined over the last decade, the most recent reports are less impressive. According to the Road Safety Annual Report 2014 released by the International Expert Network and Database on Road Safety, 2012 had the smallest annual decline in MVC-related fatality rates in ten years. In fact, these rates increased for ten countries, including Canada. This trend in MVC fatalities is supported by Transport Canada, which reports a 2.7% increase in the total number of fatalities between 2011 (2,023) and 2012 (2,077) in addition to an increase in fatalities per 100,000 population, from 5.8 in 2011 to 6.0 in 2012 (Transport Canada, 2015). MVCs remain one of the nation's leading causes of premature death and injury and remain a major population health concern for Canada (Canadian Hospital Injury Report and Prevention Program and Public Health Agency of Canada, 2013). In particular, the health and economic burden associated with MVC-related injuries weighs heavily on Canada's health care system. MVCs caused 165,172 injuries in 2012, costing the government billions of dollars, particularly for the 10,794 serious injuries requiring hospital admission (Transport Canada, 2015).

The substantial health and economic burdens associated with MVCs has prompted researchers to investigate the engineering, environmental, and human factors associated with MVCs in an effort to improve traffic safety. Generally, research and countermeasures related to the engineering and environmental factors of MVCs have shown a positive impact on traffic safety, while human factors – particularly driving behaviours – remain a complex area of MVC research and continue to contribute to the majority of MVCs (Shope, 2006).

1.1 The Role Of Human Factors In Motor Vehicle Collisions (MVCs)

MVCs are the result of the complex interactions among a multitude of factors associated with humans, vehicles, and the physical and social environments (International Traffic Safety Data and Analysis Group, 2014; Ward, Linkenbach, Keller, & Otto, 2010). There is consensus among experts in the field that human factors contribute to the majority of MVCs, more so than engineering factors (vehicle design and safety features) and factors related to the physical environment (road design and crash protective roadside objects) (Evans, 2004). Risky driving behaviours (RBDs) contribute to more MVCs than any other human factor (Evans, 1991; Evans, 1996; Peden, 2004). Such behaviours include seat belt non-compliance, cell phone-distracted driving, fatigued driving, speeding, aggressive driving, and driving under the influence of alcohol (DUIA) (Transport Canada, 2014). While much of the literature focuses on the demarcation of the human factors associated with MVC risk, less is known about the factors associated with RDBs (Iversen, 2004; Shope, 2006).

1.2 Study Rationale

Research on RDBs has identified factors such as age, sex, education, income, geographic location, and marital status as potential contributors; however, much of the literature thus far has focused on young drivers and is not generalizable to the larger driving population. Moreover, previous investigations have addressed only a limited number of RDBs or driver characteristics (Canadian Council of Motor Transport

Administrators, 2014). Inconsistent use of terminology, definitions, and measurements in the research of RDBs limits the interpretation of much of the current literature, particularly for cell phone-distracted driving and fatigued driving (Caird, Willness, Steel, & Scialfa, 2008; Di Milia et al., 2011).

Whether mental health factors contribute to the engagement in RDBs is an important area of road safety research that is fundamental to improving traffic safety. The current literature contains minimal research on mental health factors such as depression, anxiety, psychological distress, and stress and their association with risky driving (Martiniuk et al., 2010; Wickens, Smart, & Mann, 2014; Wickens et al., 2013). In addition, although research demonstrates that an individual's propensity to take risks often spans multiple behaviours, it remains unclear whether individuals who engage in RDBs also take risks in other areas of their lives (McKenna, Horswill, 2006; Turner, McClure, & Pirozzo, 2004). Factors such as riding with a drinking driver (RWDD), binge drinking, and the number of injuries sustained in the previous year may be suggestive of a general propensity towards risk-taking and contribute to the identification of factors associated with risky driving (Ivers et al., 2009; Turner, McClure, & Pirozzo, 2004). The inclusion of a comprehensive range of sociodemographic, health, and risk-related factors will address gaps in the current literature by further demarcating the factors associated with RDBs and may benefit traffic safety efforts by providing a more comprehensive profile of drivers who report RDBs.

In addition to identifying the factors associated with a broad range of RDBs, another important question is whether RDBs co-occur in the same subset of drivers. Much of the MVC research to date has investigated driver behaviours independently of one another, in combination with a restricted number of other driving behaviours, or has included a limited number of driver characteristics (Scott-Parker, Goode, & Salmon, 2015). Questions remain as to whether the same individuals or sets of individuals are commonly involved in a broad range of RDBs or, conversely, whether there are unique factors associated with different individual driving behaviours. Research to date has not investigated this "homogeneity" hypothesis in the context of RDBs, and has not employed cluster analysis to investigate a broad range of RDBs and driver characteristics. Clarifying the factors common to subsets of risky drivers may enhance prevention strategies by enabling traffic safety efforts to more accurately identify drivers who are most likely to engage in specific RDBs or patterns of RDBs.

1.3 Purpose

The purpose of this study was to fill the gaps in the literature by investigating the factors associated with risky driving using data from the Canadian Community Health Survey – 2011 (CCHS). First, descriptive statistics were used to establish the prevalence of six RDBs: seat belt non-compliance; cellphone-distracted driving; driving while fatigued; speeding; aggression; and DUIA and to determine the distribution of factors common to each driving behaviour. Second, cluster analysis of the six RDBs was employed to examine whether RDBs cluster together and reveal distinct patterns of driving behaviours. Analysis of variance (ANOVA) and Bonferroni tests were used to explore whether driver characteristics further differentiate the subgroups of drivers in the cluster solution. Finally, a series of logistic regressions verified the cluster solution by testing whether associations existed between four mental health variables and cluster membership.

1.4 Research Questions

The objectives of this study were to determine the factors associated with RDBs and to create profiles of risky drivers based on shared and distinct driving behaviours. Key research questions included:

1. What is the prevalence of the following six RDBs: seat belt non-compliance, speeding, cell phone-distracted driving, fatigued driving, aggressive driving, and alcohol-impaired driving in Canada, and what is the distribution of driver characteristics for each risky driving behaviour?
2. Do RDBs cluster together? If so, which behaviours cluster together and how does frequency of involvement in each behavior shape the resulting clusters?
3. Are there common sociodemographic and psychosocial characteristics associated with belonging to a particular risky driving cluster?

1.5 Guiding Frameworks

This study used two resources as heuristic guides to better understand the role of driving behaviour within the broader traffic system and to facilitate a structured investigation of the many factors linked to RDBs. These resources included the conceptual framework presented by Shope (2006) and the methodological strategy employed by Haddon's Matrix (Haddon, 1968; Shope, 2006).

1.5.1 Crash Risk as a Function of RDBs

Shope (2006) presents a conceptual framework to demonstrate the associations between RDBs, MVCs, injuries, and fatalities (Figure 1). The framework demonstrates how the consecutive prevention of MVC fatalities, injuries, and crashes are contingent upon the prevention of RDBs (Shope, 2006). Shope (2006) expands this framework to explore factors that may influence driving behaviour, such as personality characteristics, demographic factors, perceived environment, developmental factors, driving environment, and driving ability. This framework demonstrates the link between factors that contribute to RDBs, MVC risk, and MVC-related morbidity and mortality.

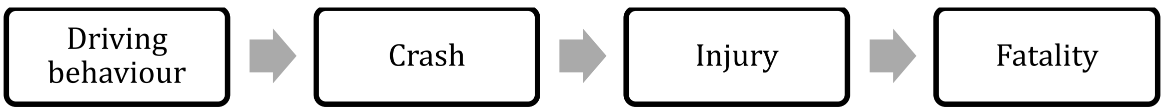

Figure 1. Crash risk as a function of risky driving behaviours
The prevention of MVC-related fatalities, injuries, and crashes ultimately requires the prevention of RDBs (Shope, 2006).

The link between RDBs and crash risk has been established in the literature (Iversen & Rundmo, 2002; Klauer et al., 2006). Particularly, there exists a considerable amount of literature devoted to the contribution of RDBs to crash risk (Aarts & van Schagen, 2006; Caird et al., 2008; Di Milia et al., 2011; Elvik, 2011; Mann et al., 2010; Transport Canada, 2014; Zador, 1991). Interpreting crash risk as a function of RDBs according to Shope's (2006) conceptual framework highlights the importance of identifying the factors associated with RDBs and is an appropriate heuristic guide for this study.

1.5.2 The Haddon Matrix

Born from William Haddon Jr.'s reductionist perspective of road safety, the Haddon Matrix is an analytic tool used to brainstorm and organize the multitude of potential factors associated with MVCs. Haddon developed the matrix in conjunction with a list of ten countermeasure strategies (not considered in this report) to more effectively reduce the losses related to humans and property (Haddon, 1968). Although some researchers believe the Haddon Matrix places too much emphasis on personal risk at the level of the individual, Haddon's work considerably broadened the previous driver-centric concept of accident prevention by considering how human, physical, and ecological factors contribute to MVC injury (Gielen & Sleet, 2003; Runyan, 2003). The Haddon Matrix is therefore an effective analytical tool to identify factors associated with MVCs from the perspective of the driver, the vehicle, and the physical and social environments.

The matrix columns represent the four major contributors to injury: (a) the host, defined as the individual at risk of injury; (b) the agent, defined as energy transferred to the host by an in inanimate (vehicle) or animate vector (assailant); (c) the physical environment defined as elements of the physical surroundings that contribute to the injury-causing event; and (d) the social environment defined as elements associated with cultural norms, politics, and legal environment. The matrix rows divide injury into three phases: pre-event, event, and post-event, with each phase allowing the opportunity for different preventative interventions. Researchers complete the matrix by identifying the potential risk and protective factors associated with each major contributor to injury at each phase of injury (Haddon, 1968). Appendix A contains an example of how the Haddon Matrix facilitates the exploration of the potential factors associated with MVCs.

When applied to RDBs, Haddon's Matrix enables the researcher to conceptualize the role of human factors, specifically RDBs, within the broader traffic system and encourages the consideration of a range of factors that contribute to engagement in RDBs. Haddon's Matrix benefits this study as it encourages brainstorming and the identification of risk factors in a systematic manner, facilitates a better understanding of the factors associated with MVCs, and considers the interaction among the various elements of the traffic system, the driver, vehicle, and the physical and social environments (Runyan, 2003). Appendix B provides an example of Haddon's Matrix as it pertains to factors potentially associated with the pre-event, event, and post-event of six RDBs.

Chapter 2: Preliminary Review Of Risky Driving Behaviours

2.0 Risky Driving Behaviours (RDBs)

Considerable research has been dedicated to the role risky driving behaviours (RDBs) play in road safety, yet there is still much to be learned about their contributing factors. Therefore, a preliminary review of the epidemiology of RDBs provides context for the exploration of the factors associated with RDBs, and a guide for the inclusion of relevant study variables in Manuscript 1 and Manuscript 2. The review includes an overview of the prevalence, impact on crash risk, contributing factors, and countermeasures or policies associated with each of the six RDBs of interest: seat belt non-compliance, cell phone distracted driving, fatigued driving, speeding, aggressive driving, and DUIA.

2.1 Seat Belt Non-Compliance

The majority of Canadian drivers are aware of the benefits of seat belt use as a protective factor for MVC-related injuries and death and all new vehicles sold in Canada are equipped with seat belts. Regardless, there is a small population of drivers who are seat belt non-compliant. Identifying non-compliant drivers is essential to reducing the number of injuries and deaths and enhancing traffic safety.

2.1.1 Prevalence of Seat Belt Non-Compliance

According to 2011 data from Transport Canada, 4.7% of Canadian drivers are seat belt non-compliant. Seat belt compliance in British Columbia, Saskatchewan, Ontario and Quebec exceed the national average, whereas seat belt compliance in all other Canadian provinces and territories, with the exception of Nunavut, is slightly below the national average (Transport Canada, 2014). A recent telephone survey of 3,888

Canadians by the Canadian Council of Motor Transportation Administrators (CCMTA) measured self-reported RDBs according to a seven-point scale ranging from one to seven (1 = *never*, 2 – 4 = *sometimes*, and 5 – 7 = *most of the time*). This study found that 16% of drivers reported seat belt non-compliance, either *sometimes* (10%) or *most of the time* (6%) (Canadian Council of Motor Transport Administrators, 2014).

2.1.2 Seat Belt Non-Compliance Impact on Crash-Related Risk

Proper seat belt use can reduce the risk of MVC fatality by 47% and MVC-related injury by 52% (Stewart, Arora, & Dalmotas, 1997). Although the impact of seat belt non-compliance on crash-related risk is difficult to measure, statistics on seat belt use among drivers involved in MVCs indirectly provides an understanding of the association between seat belt non-compliance and crash-related injury or death. Research shows the probability that a driver is seat belt compliant decreases as crash severity increases (Evans, 1996). According to Transport Canada, in 2012 32% of drivers and 37% of passengers who died in MVCs were not wearing a seat belt at the time of the crash. Similarly, among those who sustained serious injuries, 13% of drivers and 20% of passengers were seat belt non-compliant (Transport Canada, 2014).

2.1.3 Factors Associated with Seat Belt Non-Compliance

According to the literature, age is associated with seat belt use, with drivers ages 16-24 (7.0%) and 25-49 (5.2%) more likely to be seat belt non-compliant compared to drivers ages 50 and older (4.0%). This age trend for seat belt non-compliance is maintained regardless of geographic location (urban or rural) and vehicle type (pick-up trucks, cars, or mini-vans or SUVs). Male sex is also associated with seat belt use, with male drivers (5.7%) more likely to be seat belt non-compliant than female drivers (4.0%). Geography also seems to play a role in seat belt use, with lower rates of compliance among drivers in rural areas than among drivers in urban areas (Transport Canada, 2014). Other factors linked to seat belt non-compliance among drivers involved in MVCs include low level of education, unemployment, passenger status, alcohol consumption, and smoking (Canadian Council of Motor Transport Administrators, 2014; Sahai, Pitblado, Bota, & Rowe, 1998).

2.1.4 Seat Belt Non-Compliance Countermeasures

A number of seat belt policies and initiatives exist in Canada to increase seat belt compliance. Foremost, the 1991 nation-wide law requires seat belt and child restraint use in all vehicle seats. Violators of the law are fined between $75 and $250 and may be subject to the demerit of one to four points from their driving record. Other successful efforts to decrease seat belt non-compliance include the Periodic Selective Traffic Enforcement Programs (STEPs) and feedback signs on the sides of roadways to inform drivers of the rate of seat belt use on a daily basis.

More recently, the National Occupant Restraint Program (NORP 2010) was established as a part of Road Safety Vision 2010 to increase seat belt compliance. The NORP 2010 aimed to achieve a 95% seat belt compliance rate by occupants of all types of vehicles by 2010 and to reduce seat belt non-compliance-related morbidity and mortality. The program included multiple surveys in rural and urban areas of Canada conducted by Transport Canada to determine the rate of seat belt compliance, expansion of public education regarding the importance of proper child vehicle safety and booster seat use, increased penalties for lack of seat belt use and elimination of seat belt exemptions for certain passengers. Although NORP 2010 was successful in achieving an overall national seat belt compliance rate of 95%, some areas of Canada fell short of this goal (Transport Canada, 2014).

2.2 Distracted Driving

Distracted driving is defined by the CCMTA as, "the diversion of attention from driving, as a result of the driver focusing on a non-driving object, activity, event or person" (CCMTA - STRID Sub-group on Distraction). According to Transport Canada data from 2003-2007, approximately 11% of MVC-related injuries and fatalities involve driver-distraction as a factor. Such distractions include cell phone use, eating or drinking, talking to passengers, smoking and adjusting radio controls (CCMTA - STRID Sub-group on Distraction). Other research based on self-reported data report the prevalence of cell phone-distracted driving to be 30% (Canadian Council of Motor Transport Administrators, 2014).

2.2.1 Prevalence of Cell Phone-Distracted Driving

Research pertaining to cell phone use as a form of distracted driving is becoming more common due to the increase in the number of cell phone users and their propensity to use their cell phone while driving (Strayer, Drews, & Johnston, 2003; World Health Organization, 2011). The prevalence of cell phone-distracted driving is unclear. In a survey of Canadian drivers by Vanlaar et al. (2006), 37% of drivers reported cell phone-distracted driving in the previous seven days. The CCMTA reports a slightly lower prevalence of 30%, while a survey of 622 Albertans found 52% of respondents reported cell phone use while driving, 45% of whom reported using hands-free devices (Canadian Council of Motor Transport Administrators, 2014; Nurullah, Thomas, & Vakilian, 2013).

2.2.2 Cell Phone-Distracted Driving Impact on Crash Risk

Cell phone use (hand-held and hands-free) while driving reduces awareness of visual input and has detrimental effects on decision-making and performance while driving, which in turn, may increase the risk of driver errors, near-crashes, or crashes (Harbluk, Noy, Trbovich, & Eizenman, 2007; CCMTA - STRID Sub-group on Distraction). Although many studies show associations between cell phone use while driving and an increase in the risk of an MVC, study quality to date is poor and few have proven causality (Caird et al., 2008; McEvoy et al., 2005; Redelmeier & Tibshirani, 1997; Törnros & Bolling, 2006). A meta-analysis by Elvik (2011) indicates that drivers who engage in cell phone-distracted driving are 2.86 times more likely to incur an MVC than drivers who do not use their cell phone while driving. According to research by Redelmeier, D.A. & Tibshirani, R.J. (1997), the relative risk of MVC is four times higher for cell phone-distracted drivers compared to those who refrain from using their cell phone while driving (RR = 4.3, 95% CI 3.0-6.5, $p<.05$). McEvoy et al. (2005) demonstrated comparable results (OR = 4.1, 95% CI: 2.2-7.7, $p<0.001$), which also demonstrated increased risk of MVC for both handheld and hands-free cell phones (OR = 4.9, 95% CI: 1.6-15.5, $p = 0.003$ and OR = 3.8, 95% CI: 1.8-8.0, $p<.001$, respectively).

Hands-free cell phones, although developed with the intention to reduce the level of physical distraction for drivers, reduce the driver's visual awareness of traffic and vehicle instruments, subsequently reducing the driver's overall control of the vehicle (McCartt, Hellinga, & Bratiman, 2006). Such cognitive distraction has a very real impact on driver attention, supporting the notion that hands-free cell phones are no safer than handheld cell phone use while driving (Brace et al., 2007; Burns, Lécuyer, & Chouinard, 2008; Ishigami & Klein, 2009; McCartt et al., 2006; Svensen & Patten, 2005; World Health Organization, 2011). Many drivers consider hands-free cell phone use as a safe alternative to hand-held cell phone use while driving (White, Eiser, & Harris, 2004; Zhou, Wu, Rau, & Zhang, 2009). In addition, due to variation in driver attitudes, perceptions of risk, and motivations to use hands-free devices, hands-free cell phone-distracted driving is excluded from the present study (White, Eiser, & Harris, 2004; Zhou, Wu, Rau, &Zhang, 2009).

2.2.3 Factors Associated with Cell Phone-Distracted Driving

The majority of research shows an inverse relationship between age and cell phone-distracted driving. Interestingly, younger and less experienced drivers may be more susceptible to the detrimental effects of cell phone-distracted driving and text messaging (World Health Organization, 2011). This may be due to the cognitive changes teenagers and young adults experience that increase the likelihood of distraction (Brace, Young, & Regan, 2007). Recent research indicates that middle-aged drivers also engage in cell phone-distracted driving (Asbridge, Brubacher, & Chan, 2013).

Results of research examining sex as a potential contributing factor to cell phone-distracted driving are mixed. The majority of the literature indicates that males are more likely than females to talk on a cell phone and text message while driving (Hancock, Lesch, & Simmons, 2003; World Health Organization, 2011). The literature also indicates links between low level of education and urban location and cell phone-distracted driving, but further research is needed to confirm these findings (Asbridge et al., 2013; Caird et al., 2008; Hancock et al., 2003; McEvoy et al., 2005; Redelmeier & Tibshirani, 1997; World Health Organization, 2011).

2.2.4 Cell Phone-Distracted Driving Countermeasures

Despite the challenges associated with research, a number of literature reviews have provided sufficient cumulative evidence of the dangers associated with cell phone-distracted driving to encourage safe driving practices and policy change (Caird et al., 2008; Ishigami & Klein, 2009; Sugano, 2005). As of January 2012, all provinces prohibit hand-held devices, with only three provinces specifying the prohibition on hands-free devices for novice drivers (British Columbia, Saskatchewan, and the Yukon Territory) (Current Legislation - Transport Canada, 2012). Penalties associated with this law differ by province and include fines ranging from $100 to $400.

Additional methodologically sound research on cell phone-distracted driving is imperative to clarify its prevalence, increase awareness of the dangers associated with this RDB, and encourage future motor vehicle policy amendments (Caird et al., 2008; Nelson, Atchley, & Little, 2009; Patel, Ball, & Jones, 2008; Vanlaar & Yannis, 2006).

2.3 Fatigued Driving

The literature contains very little research pertaining to fatigued driving, mostly due to the difficulties in measurement and variations in terminology of fatigued driving and research lacking methodological rigor (Di Milia et al., 2011). This may be due, in part, to the difficulties associated with measurement and the underestimation of fatigue-related MVC by police reports resulting from inconsistent use of a definition of fatigue in the documentation of such collisions. Fatigue-related MVCs are commonly reported as 'loss of vehicle control' or 'inattention' (Vanlaar, Simpson, Mayhew, & Robertson, 2008). Driver recall also plays a role in how a fatigue-related MVC is reported. Many drivers may provide inaccurate recall of the events leading up to a crash due to the traumatic effects and/or arousal post-collision or data may be unattainable due to driver and/or passenger fatality (Radun & Radun, 2009). The importance of using a standard definition of driver fatigue is therefore imperative to accurate estimates of the prevalence of fatigue-related MVC.

2.3.1 Prevalence of Fatigued Driving

Millions of Canadians drive while fatigued each year. According to Vanlaar, Simpson, Mayhew, & Robertson (2008), almost 60% (4.8 to 5. 5 million drivers) of Canadian drivers report occasionally driving when fatigued, 15% (1 to 1.5 million) have fallen asleep or nodded off while driving, and 2% (97,000 to 29,000 drivers) of drivers report a fatigue-related MVC in the past 12 months (CCMTA - STRID Sub-group on Fatigue). Survey data from 2014 suggest that the prevalence of fatigued driving is even higher, at 72% overall, with 61% of drivers reporting *sometimes* and 11% *most of the time* (Canadian Council of Motor Transport Administrators, 2014).

2.3.2 Fatigued Driving Impact on Crash Risk

Driver fatigue results in longer reaction times and unstable operation of the vehicle, the magnitude of which increases over time. Subsequently, the most common errors made by fatigued drivers are following vehicles too closely and loss of control of the vehicle (Elzohairy, 2008). In fact, similar to cell phone use while driving, the impact of driver fatigue is comparable to driving under the influence of alcohol (Williamson, Feyer, Mattick, Friswell, & Finlay-Brown, 2001). According to Transport Canada, driving performance at 19 hours without sleep is similar to that of an individual driving with a BAC of 0.05 (CCMTA - STRID Sub-group on Fatigue).

2.3.3 Factors Associated with Fatigued Driving

The literature contains very little empirical research of the factors associated with fatigued driving. Factors linked to self-reported fatigued driving include young age and higher income (Canadian Council of Motor Transport Administrators, 2014). A review of the literature by Di Milia, et al. (2011) also suggests links with young age in addition to male sex. Other factors that may impact the risk of a fatigue-related MVC include individuals with sleep disorders such as narcolepsy, alcohol use, medications that cause drowsiness, trip duration, certain occupations (such as commercial vehicle operators required to drive for long periods of time) and night or rotating shift workers (Vanlaar, Simpson, Mayhew, & Robertson, 2008b).

2.3.4 Fatigued Driving Countermeasures

Apart from long haul truck drivers, no specific law pertaining to fatigue-related MVCs exists in Canada. Effective measures implemented to reduce fatigue-related MVC include highway rumble shoulders and center-line rumble strips. Rumble shoulders alert drivers who may be drifting off the road and reduce the number of single-vehicle crashes by 18-21%. Center-line rumble strips similarly alert drivers who may begin to cross over into oncoming traffic and have been shown to reduce the number of head-on crashes by 25% ('Fatigue Impairment', 2008). Other efforts to reduce fatigue-related MVC include vehicle devices to detect eye lid closure, head nodding and other indications of driver fatigue, driver reviver stations to enable fatigued drivers to pull over and rest and drowsy driver signs along roadways to remind drivers of the risks associated with fatigued driving. However, the effectiveness of these methods to reduce fatigue-related MVC remains unknown ('Fatigue Impairment', 2008).

2.4 Speeding

Speeding is a major problem among Canadian drivers and is a behaviour that endangers drivers, passengers, pedestrians, cyclists, and other road users on a daily basis. Vehicles going off the road, hitting objects or persons, or head-on collisions are common occurrences in speeding related single car crashes (Transport Canada, 2013). Despite speed limits, penalties, and the well-known risks associated with speeding, many drivers continue to drive at unsafe speeds.

2.4.1 Prevalence of Speeding Drivers

In Canada, 70% of drivers admit to sometimes exceeding the speed limit. Speeding is a factor in 19-20% of MVC-related injuries and in 25-27% of MVC-related fatalities; as a reason for MVC fatalities, it ranks second only to impaired driving (Quimby, Maycock, Palmer, & Buttress, 1999; Transport Canada, 2014; Vanlaar, Robertson, & Marcoux, 2008; Vingilis & Wilk, 2010).

2.4.2 Speeding Impact on Crash Risk

It is well known that speeding increases the risk of an MVC due to the increased braking and reaction distances required for manoeuvres at high speeds (Nilsson, 2004; Wilson et al., 2010). Research shows the odds of incurring an MVC-related injury are over two times greater for individuals who sometimes, rarely, or never obey the speed limit compared to those who mostly or always obey speed limits (Vingilis & Wilk, 2010). The magnitude of the impact of speeding on MVC risk varies, depending on characteristics of the road and the driver, as well as driving behaviours, but research shows that a small increase in speed results in an exponential increase in the risk of an

MVC regardless of road type, particularly for single car collisions (Aarts & van Schagen, 2006; Kloeden et al., 1997). Similarly, a small reduction in speed can greatly reduce the risk of MVC. The relationship between MVC-related fatalities, serious injuries, and speed is illustrated by Nilsson's (2004) Power Model, which shows that a five percent decrease in average speed results in approximately a 10% decrease in MVC injuries and a 20% decrease in MVC fatalities.

2.4.3 Factors Associated with Speeding

Research has shown age and sex as important factors associated with speeding. In Canada, drivers under the age of 30 are more likely than older drivers to report non-adherence to speed limits. In fact, drivers under the age of 45 are responsible for 80% of MVC speeding-related injuries and fatalities and Canadian drivers aged 16-24 represented almost half of all speeding-related MVCs on urban roads between 2002-2004 (Transport Canada, 2013). Research also shows that young male drivers are more likely to speed than older drivers of either gender (Transport Canada, 2013). According to Vingilis and Wilk, (2010) Canadian males – particularly young males – are less likely to adhere to speed limits than females (Vingilis & Wilk, 2010). Likewise, Rhodes and Pivik (2011) found a similar trend for sex in speeding-related MVCs as well as speeding-related fatality risk. In addition to age and sex, the presence of peer passengers may also be a factor associated with speeding. A study by Simons-Morton et al. (2005) demonstrated that young-adult drivers who drove with peer-passengers were more likely to drive faster, speed, tailgate and engage in other RDBs. Therefore, it is not surprising that between 2002 and 2004, 80% of passengers in fatal crashes were in a vehicle with a speeding driver of comparable age (Transport Canada, 2013). Other factors linked to speeding are alcohol and urban location. Vingilis et al. (2007) estimate that alcohol is a factor in one-third of all speeding-related fatalities, and according to Transport Canada, MVCs involving speeding, alcohol, and young drivers most commonly occur on urban roads at night (Transport Canada, 2013).

2.4.4 Speeding-Related Countermeasures

Countermeasures related to speeding include the implementation of speed radar cameras that record a speeding driver's license plate and generate a summons that is sent to the address of the vehicle owner. Such municipal speed radar cameras programs exist in Edmonton and Calgary, Alberta and in Winnipeg, Manitoba. Preliminary studies show a 44% decrease in speeding and 94% decrease in extreme speeding (driving 50 km over the speed limit) (Parliament of Canada, 2005). Additionally, amendments to laws prohibiting street racing (driving 50 km over the speed limit) in Ontario increased penalties to perpetrators, including automatic license suspension, vehicle impoundment, or fines up to $10,000. The amendments reduced speeding by 50% and prompted the implementation of similar anti-street racing laws in British Columbia in 2010 and Quebec in 2011 (Parliament of Canada, 2005; Vingilis & Wilk, 2010). Finally, countermeasures specific to younger drivers include amendments to the graduated licensing program, which increased penalties for novice drivers caught speeding through license suspension of first-time ticketed drivers (Vingilis & Wilk, 2010).

2.5 Aggressive Driving

Driver aggression is defined as, "attempts by drivers to direct aggression toward other road users, including attempts to injure or damage" (Linden et al., 2010). However, driver aggression-related MVC research often uses a more liberal interpretation of aggressive driving and includes RDBs that neglect safety such as speeding, street racing, tailgating, rapid acceleration, weaving in and out of traffic, negligence of traffic control devices and thrill seeking (Paleti et al., 2010; Vanlaar, Simpson, Mayhew, & Robertson, 2007). More recently, the term *road rage* has been used to describe drivers who exhibit extreme aggressive driving behaviour. Road rage is defined as a "driver or passenger who attempts to kill, injure, or intimidate a pedestrian or another driver or passenger or to damage their vehicle in a traffic incident" (Smart & Mann, 2002). Although the definition of driver aggression and road rage are similar and at times overlap, the literature pertaining to both driver aggression and road rage will be considered in aggressive driving-related MVCs.

2.5.1 Prevalence of Aggressive Driving

Driver aggression is not foreign to Canadian drivers, but estimates of the prevalence of aggressive driving are variable due to the paucity of empirical research. Almost 47% of Canadian drivers have reported being a victim, at some point in time, of mild forms of driver aggression and 7.2% have reported experiencing severe driver aggression, while 31.7% and 2.1% of drivers reported committing mild and severe forms of driver aggression, respectively (Smart & Mann, 2002a; Smart & Mann, 2002). Similarly, self-reported data from 2014 showed that 39% of drivers admitted to aggressive driving (Canadian Council of Motor Transport Administrators, 2014).

2.5.2 Aggressive Driving Impact on Crash Risk

There is minimal research of the impact that aggressive driving has on crash risk. Beirness et al. (2001) estimates that aggressive driving is responsible for up to 18% of

MVC-related fatalities and injuries. Research by Mann et al. (2007) demonstrated an increased risk of MVC among road rage victims and perpetrators. Results showed that victims were 2.84 times more likely to have reported an MVC and perpetrators of road rage were 2.24 times more likely to have reported an MVC compared to non-victimized drivers and road rage non-perpetrators, respectively.

2.5.3 Factors Associated with Aggressive Driving

The majority of research shows that younger drivers are more likely to engage in aggressive driving than older drivers, however some research suggests that aggressive driving may also be a factor among middle-aged drivers (Vanlaar, Simpson, Mayhew, & Robertson, 2007). Asbridge, Smart, and Mann (2003) demonstrated an inverse association between road rage and age, aside from a higher probability of both road rage victimization and offending among drivers between the ages of 50 and 64. Looking at sex, although males and females were equally likely to be victims of road rage, male drivers are twice as likely as females to exhibit road rage (Asbridge, Smart, & Mann, 2003; Vanlaar et al., 2007). Previous road rage victimization is also linked to aggressive driving, with road rage offenders five times more likely to have experienced road rage as a victim than non-offenders. Other sociodemographic factors are linked to aggressive driving, including single marital status, urban residence, and high-income and educational attainment (Asbridge et al., 2003; Canadian Council of Motor Transport Administrators, 2014).

Research also identifies a number of health factors that contribute to aggressive driving. Stress and driver aggression are closely associated, as driver aggression is one of the most common reactions for drivers experiencing stress, particularly during rush hour traffic (Hennessy & Wiesenthal, 1997; Matthews et al., 1998). In addition, aggressive drivers are more likely to report psychological distress, problems with alcohol, and illicit drug use (Butters, Mann, & Smart, 2006; Mann, Smart, Stoduto, Adlaf, & Ialomiteanu, 2004; Smart, Asbridge, Mann, & Adlaf, 2003).

2.5.4 Aggressive Driving Countermeasures

The countermeasures or policy reforms to reduce aggressive driving are minimal aside from campaigns to promote courteous and safe driving and those that are speeding related. For example, many municipalities in Canada have implemented red light cameras to photograph vehicles that violate traffic signals as a means to reduce aggressive driving at intersections. A study in the United States showed that this method reduced fatal red-light collisions in US cities by 35% between 2004 and 2008 (Transport Canada, 2014).

2.6 Driving Under The Influence Of Alcohol (DUIA)

The effects of alcohol on driving ability are well documented. Foremost, alcohol depresses the area of higher motor functions, reaction time and vision, and eventually affects balance, sensory perception, and coordination (Moskowitz & Fiorentino, 2000). Such impairments are implicitly detrimental to driving ability and significantly increase the risk of MVCs (Mann et al., 2010). BAC is measured by volume in terms of the number of milligrams of alcohol per milliliter of blood (mg/ml) and is often expressed as a percentage, whereby a BAC of 80 mg of alcohol per 100 ml of blood means that 0.08% of a person's blood by volume is alcohol. The current legal BAC limit in Canada is 80 mg of alcohol per 100 ml of blood or 0.08% (Transport Canada, 2008).

2.6.1 Prevalence of DUIA

Recent research indicates an increase in the prevalence of self-reported DUIA (Vanlaar et al., 2012). Using the National Opinion Poll on Drinking and Driving, Vanlaar, et al. (2012) demonstrated a significant increase in the percentage of Canadian drivers who reported driving after drinking any amount of alcohol in the past 30 days from 14.7% in 2005 to 19.2% in 2011 ($Z = -2.88$; $p = 0.004$). The most recent self-reported data (2014) indicates an increase in the prevalence of self-reported DUIA, with 24% of drivers reporting either *sometimes* (21%) or *always* (3%) driving within two hours of consuming two or more alcoholic drinks (Traffic Injury Research Foundation, 2014).

2.6.2 DUIA Impact on Crash Risk

Research demonstrates an exponential increase in the relative risk of fatal and non-fatal MVCs as BAC increases (Blomberg, Peck, Moskowitz, Burns, & Fiorentino, 2009; Borkenstein, 1974; Peck, Gebers, Voas, & Romano, 2008). Drivers with a BAC between 0.05-0.09% are nine times more likely to incur an MVC than a driver with zero BAC (Zador, 1991). Although the majority of alcohol-related MVC fatalities involve BACs over 0.08%, many DUIA-related fatal MVCs occur among drivers with BAC below that of the legal limit (Phillips & Brewer, 2011; Transport Canada, 2008; Transport Research International Documentation (TRID), 2013). The impact of DUIA on crash risk is highlighted by 2010 data showing that of all drivers killed in MVCs in Canada who were tested for alcohol, 33.6% of drivers had positive BAC (Traffic Injury Research Foundation, 2014).

2.6.3 Factors Associated with DUIA

A number of the factors are associated with DUIA and alcohol-related crashes. An inverse relationship exists between age and the risk of DUIA-related MVC, where younger drivers are at an increased risk compared to older drivers (Blomberg et al., 2009; Zador, 1991). There is also a distinct gender difference in alcohol-related MVC injuries and fatalities, with males accounting for 85.4% of Canadian drivers involved in a fatal alcohol-related MVC (Transport Research International Documentation (TRID), 2013; Zador, 1991).

In addition to the aforementioned factors that directly affect driving ability and increase the risk of MVC, factors associated with alcohol consumption also increase the odds of a MVC. The likelihood of MVC is higher among those who engage in binge drinking, particularly among youth. In an investigation of the effects of health and substance use on future MVC injuries among over 16,000 Canadian drivers, Vingilis and Wilk (2008) used path analysis technique and found Canadian youth (under 30 years of age) who engage in binge drinking are 2.3 times more likely to have an MVC than those who do not binge drink. Alcoholism also increases the risk of MVC, whereby alcoholics incur almost twice the risk of an MVC and related mortality compared to that of the general population (Macdonald, Anglin-Bodrug, Mann, & Chipman, 2005). A study conducted in Ontario, Canada by Mann, Stoduto, Vingilis and Asbridge, et al., (2010) demonstrates that individuals who reported alcohol dependence or alcohol-related problems were more likely to have an MVC than those who did not report such behaviours. This evidence highlights the importance of the relationships among the cognitive and physiological effects of alcohol, the unique personality characteristics of individuals with drinking problems, and increased risk of MVCs (Mann et al., 2010).

2.6.4 DUIA Countermeasures

Canada's initial DUIA-related federal policy, Section 253 of the Criminal Code of Canada (1968/1969), prohibits driving a motor vehicle with a BAC over 0.08% as determined by a breathalyzer test. The minimum penalties associated with DUIA in Canada increase in severity each time a driver is charged. The first offence results in a $1,000 fine and a one-year driving prohibition, the second offence results in 30 days in jail and two years of driving prohibition and the third and subsequent offences thereafter result in 120 days in jail and three years of driving probation. According to Asbridge, Mann, Flam-Zalcman, and Stoduto (2004), BAC-related policies such as the Breathalyzer Law are effective in reducing the prevalence of alcohol-related BACs. Asbridge et al. (2004) found that the breathalyzer law was effective in reducing alcohol-related MVC fatalities in Ontario by 18% and had a long-term effect on the prevalence of alcohol related MVC fatalities.

The implementation of DUIA legislation such as the breathalyzer law decreased prevalence of DUIA in the 1970s and 1980s. However, progress stagnated in the 1990s and 2000s and led to the development of various initiatives at the provincial level. For example, the ignition interlock system is employed in all provinces in an attempt to further reduce DUIA. The ignition interlock system requires DUIA offenders to provide a breath sample prior to starting the car. If the breath sample contains a BAC above a court-determined limit, the vehicle will not start. The mandatory ignition interlock systems for drivers convicted of DUIA have shown to reduce re-arrest rates for first and repeat offenders, but the effect of these systems on the prevalence of DUIA once the device is removed is still unknown (Elder et al., 2011; Willis, Lybrand, & Bellamy, 2004). Furthermore, the mandatory ignition interlock system lacks a national standard, resulting in inconsistency in the technical standards across Canada. The creation of an approved standard would benefit future use of a mandatory ignition interlock system for DUIA convicted drivers and research to determine its effectiveness (Beirness & Boase 2007).

The establishment of anti-DUIA campaigns have also decreased DUIA in Canada. The Canadian Council of Motor Transport Administrators' Strategy to Reduce Impaired Driving (STRID) promotes awareness, encourages police enforcement, and advocates for treatment programs to address impaired driving. A major issue addressed by STRID is the implementation of a policy at the national level to address DUIA at BAC levels under the legal limit (CCMTA - Strategy to Reduce Impaired Driving (STRID) 2010). Similarly, the establishment of Mothers Against Drunk Driving (MADD) in the 1980's and its emphasis on driver awareness and education around drinking and driving has had a profound effect on the reduction in the number of alcohol- and non-alcohol-related fatalities nation-wide (Asbridge et al., 2004).

2.7 Summary

The preliminary review of the literature provides a number of key findings pertaining to the prevalence, impact on crash risk, contributing factors, and countermeasures associated with each of the six RDBs of interest: seat belt non-compliance, cell phone distracted driving, fatigued driving, speeding, aggressive driving, and DUIA. First, sex and age are important sociodemographic factors associated with drivers who engage in RDBs, however the demarcation of the factors associated with RDBs remains an important area of research that requires further exploration. Second, although the literature indicates that factors associated with mental health play a role in aggressive driving, there is a lack of research investigating their contribution to other RDBs. Third, compared to the evidence on seat belt non-compliance, DUIA, speeding, and aggressive driving, less is known about cell phone-distracted driving and fatigued driving. Part of the difficulty is due to variations in terminology and methodological challenges associated with the measurement of these RDBs. Fourth, although further research is required to establish a causal relationship for some RDBs, it is clear that RDBs are linked to crash risk. Fifth, although the majority of Canadian drivers are seat belt compliant and do not engage in DUIA, many drivers engage in speeding, fatigued driving, cell phone-distracted driving, and aggressive driving, despite the dangers associated with these driving behaviours. Finally, traffic safety efforts have implemented many countermeasures and policies to reduce the prevalence of drivers who engage in RDBs. Although many of these efforts have been successful, those that influence a change in traffic safety culture, such as those for seat belt non-compliance and DUIA, seem to be the most successful.

Overall, the preliminary review of the prevalence, impact on crash risk, contributing factors, and countermeasures or policies associated with each of the six RDBs indicates that much of the research thus far is restricted to subpopulations of drivers, or investigates a limited number of RDBs (expanded upon in Manuscripts 1 and 2). The use of population-based data will enable a more comprehensive profile of drivers who engage in RDBs.

Chapter 3: Cluster Analysis

3.0 Cluster Analysis

Cluster analysis, also known as data segmentation, is a descriptive, unsupervised, and noninferential procedure useful for exploratory data analyses of large data sets (Hastie, Tibshirani, & Friedman, 2009). The purpose of cluster analysis is to explore the structure of the data; it is not measurement of a specific target variable. Cluster analysis sorts the data into clusters that share similar characteristics, by grouping objects (observations, events, or cases) into mutually exclusive clusters that have similar profiles according to the attributes (variables) associated with each object (Bailey, 1994). An ideal cluster solution is one that contains clusters with a high level of similarity in regards to the characteristics they possess (maximizing the within-cluster homogeneity) and a high level of dissimilarity from cases in other clusters (maximizing the between-cluster heterogeneity) (Bible, Datta, & Datta, 2013). In this study, the objective was to sort the data according to drivers' responses to the six RDBs of interest (the clustering variables).

Cluster analysis is beneficial to researchers working with large data sets containing numerous observations, as it compresses the data into manageable groups that share similar attributes. The cluster solution's simplified groups may reveal patterns in the data or associations between variables that were otherwise hidden in large data sets (Everitt, Landau, Leese, & Stahl, 2011). Data reduction allows researchers to identify unique subgroups or patterns within the data, making it a useful tool for taxonomy description (Bailey, 1994). Data reduction and taxonomy of large data sets may also act as starting points for subsequent analyses, enabling the generation of new hypotheses or further investigation of previously defined hypotheses (Gan, Ma, & Wu, 2007). Cluster analysis methods have been employed in many fields of research such as the social sciences, biology, statistics, data mining, marketing, and information machine learning. For example, employment of clustering methods can group documents for web browsing, group consumers according to brand loyalty and price consciousness, or determine genes and proteins with similar functions.

Experts in the field of cluster analysis have devised a vast number of clustering methods algorithms. Despite the multitude of methodologies available, all clustering approaches have the same purpose, to unveil natural structure, patterns, or latent groups within the data by using a measure of proximity (similarity or dissimilarity) to create subgroups of objects (cases) with similar profiles. All clustering methodologies aim to maximize within-cluster homogeneity and between-cluster heterogeneity; in other words, clustering minimizes intra-cluster distance and maximizes inter-cluster distance (Figure 2) (Everitt et al., 2011).

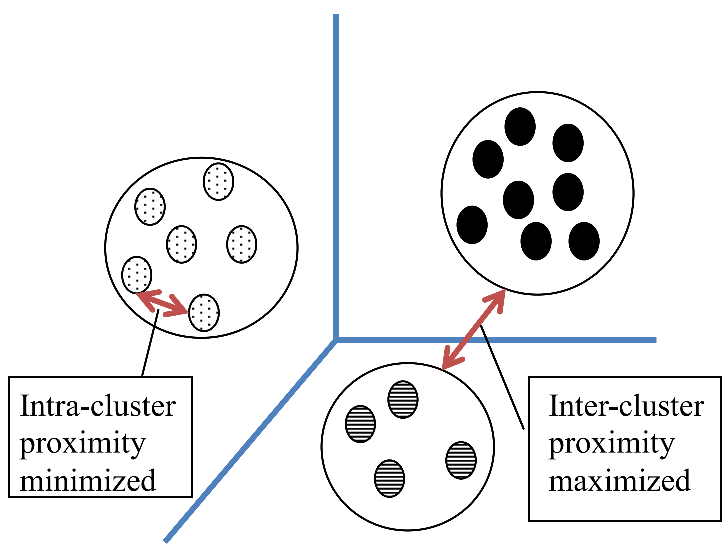

Figure 2. Within-cluster homogeneity and between-cluster heterogeneity

The two most frequently employed methods of cluster analysis are hierarchical clustering (connectivity models) and partition clustering (k-means or k-median cluster analysis) (Gan et al., 2007) (Figure 3). To examine RDBs among Canadian drivers, this study employed Ward's hierarchical agglomerative clustering method on a subsample of data (using the six RDBs as grouping variables) to determine the ideal number of clusters for a *k*-means cluster analysis on the full sample. Therefore, the following sections will focus on these two methods of clustering. For a more comprehensive explanation of the vast number of clustering methods and algorithms, refer to Everitt et al. (2011) and to Bailey (1994).

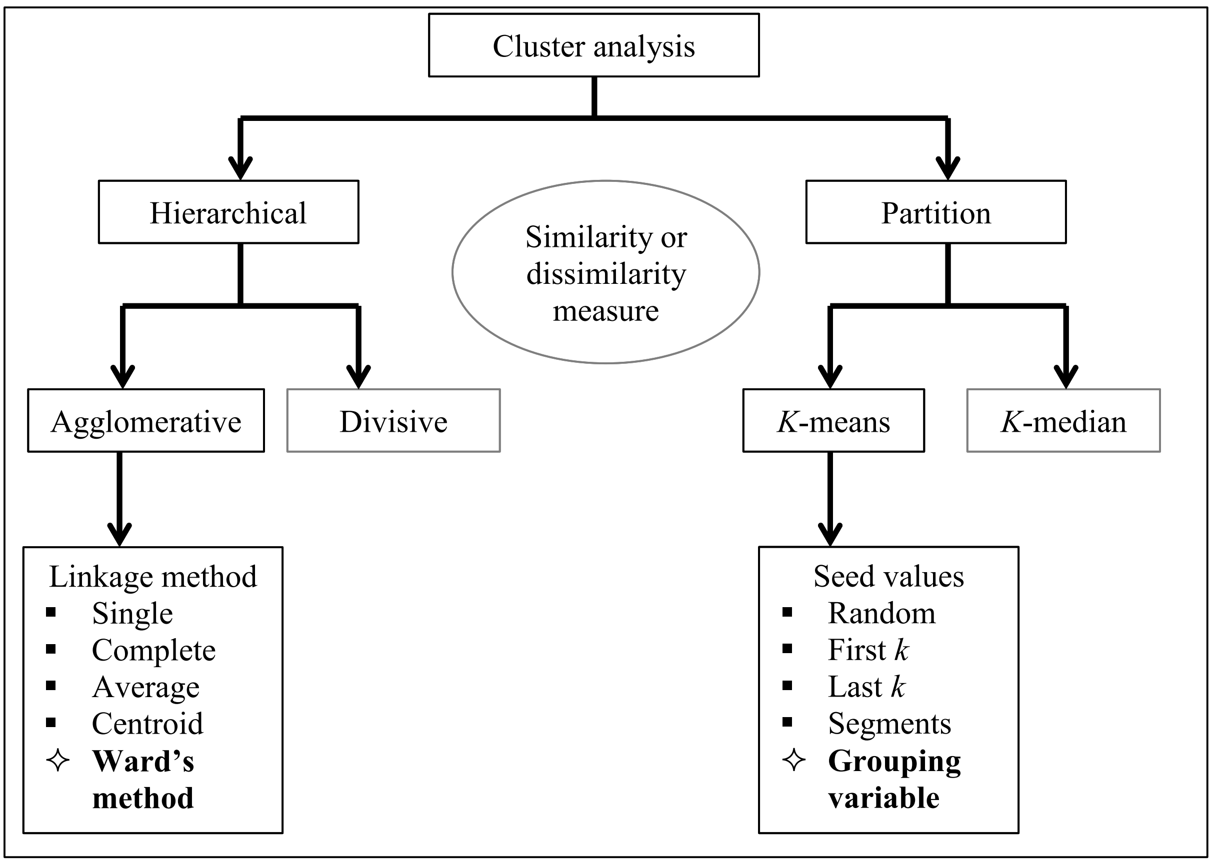

Figure 3. Cluster analysis methodologies

3.1 Hierarchical Cluster Analysis

Hierarchical cluster analysis is a stepwise iterative process that uses a measure of proximity to merge or divide clusters to produce a set of nested clusters. Each mutually exclusive cluster contains objects that are highly similar to one another and distinctly different from objects in other clusters. Hierarchical clustering can be either agglomerative or divisive (Figure 4), both of which produce a hierarchical structure called a dendrogram that manifests the order in which clusters merge or divide (Figure 5) (Abbas, 2008; Everitt et al., 2011). Divisive cluster analysis begins with all objects in one large cluster, which iteratively divides into sub-clusters of similar objects based on a measure of proximity. There are very few hierarchical divisive methodologies in the literature, and many statistical analysis software programs do not provide commands for such procedures, therefore, discussion of hierarchical divisive methods of clustering will conclude here.

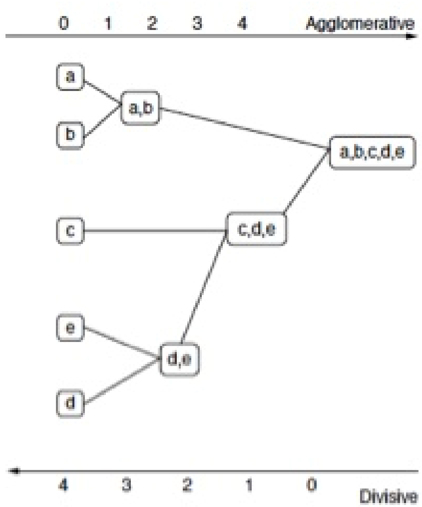

Figure 4. Agglomerative and divisive hierarchical cluster analysis
Reprinted from Cluster Analysis, 5th Edition (page 72), by B.S. Everitt, S. Landau, M. Leese, & D. Stahl, 2011, John Wiley & Sons, Ltd., 5th.

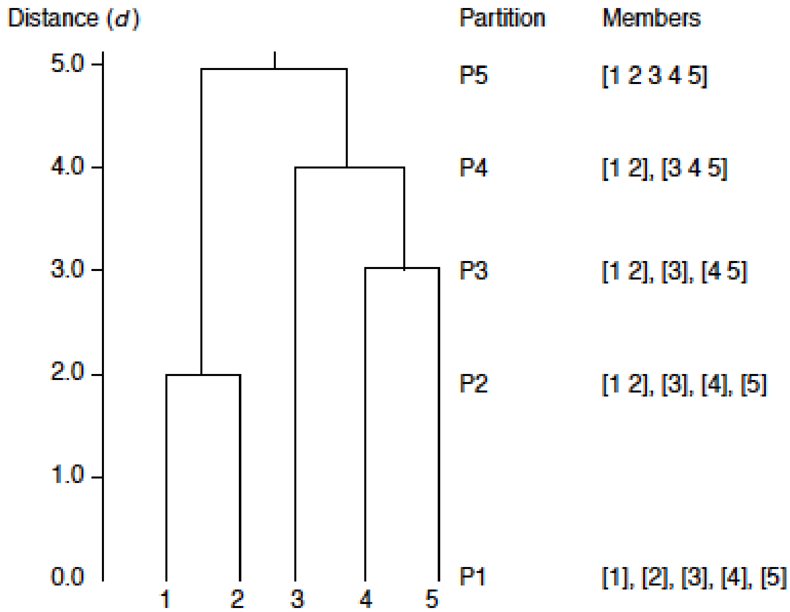

Figure 5. Example of a dendrogram generated by hierarchical agglomerative cluster analysis
Reprinted from Cluster Analysis, 5th Edition (page 75), by B.S. Everitt, S. Landau, M. Leese, & D. Stahl, 2011, John Wiley & Sons, Ltd., 5th.

Hierarchical agglomerative clustering groups objects (measures, events, samples, or patterns) according to their similarity on attributes (variables or measures). Conceptually, the objects of interest are points or vectors in a multi-dimensional space, where each dimension represents an attribute. The data matrix represents the objects and attributes as an *m* by *n* matrix. The matrix of *m* rows of objects and *n* columns of attributes are the basis for the calculation of a proximity matrix, the initial step in the clustering procedure for most hierarchical clustering methods (Bailey, 1994).

This method of cluster analysis is termed "agglomerative" because the creation of new clusters combines previously merged clusters (nested clusters) (Abbas, 2008). The clustering process for some hierarchical agglomerative methods begins with the calculation of a proximity matrix and each object as a cluster of one (singleton cluster). The two singleton clusters that are most similar to one another are merged to form a new cluster and the proximity matrix is updated. One merger between two clusters occurs per iteration of the clustering process. The iterative process continues to merge the most similar clusters and update the proximity matrix until there exists one large cluster. The basic algorithm for hierarchical agglomerative clustering is (Tan, Steinbach, & Kumar, 2006):

1. Compute the proximity matrix.
2. Merge the closest two clusters.
3. Update the proximity matrix
4. Repeat steps 2 and 3 until only one cluster remains.

3.1.1 Measure of Proximity

The proximity measure is an important component of cluster analysis, as all cluster analysis algorithms are based on an index of similarity or dissimilarity of data points (Jain & Dubes, 1988). Quantitative description of the similarity (or dissimilarity) of two data points or clusters requires a measure of proximity in the form of similarity measures, similarity coefficients, dissimilarity measures, or distances (Gan et al., 2007).

To date, it is unclear which proximity measure is optimal for cluster analysis (Everitt et al., 2011). Comparative studies by Cheetham and Hazel (1969), Boyce (1969), and Williams et al. (1966) show that of the many proximity measures applicable to clustering methods, there is no definitive optimal measurement. Researchers must choose a measure of proximity that suits their research purposes and is appropriate for the clustering method employed. One of the most frequently used measures of proximity (and the measure of proximity used in this study) is the squared Euclidean distance (Everitt et al., 2011). The squared Euclidean distance is a measure of dissimilarity (the sum of the squared differences between the values of two attributes) between two data points *x* and *y* in multidimensional (*d*) space and is defined as

$$d_{seuc}(x,y) = \sum_{j=1}^{d}(x_j - y_j)^2 ,$$

where x_j and y_j are the values of the j^{th} attribute of *x* and *y*, respectively. Smaller squared Euclidean distance measures indicate a higher degree of similarity between two objects than larger measures.

3.1.2 The Proximity Matrix

A proximity matrix is an *m* by *m* matrix of all the pairwise indices of proximity (similarity or dissimilarity) of a data set (Figure 6). The calculation of a proximity matrix determines which two objects (or clusters) are most similar and merged at each step of the clustering process (Bailey, 1994). During each iteration of the clustering process, the measure of proximity for each pair of objects or singleton clusters is calculated (Figure 7a), the most similar objects or clusters are merged (Figure 7b), and the proximity matrix is updated and reduced by one (the two singleton clusters are replaced with the merged cluster) (Figure 7c) (Everitt et al., 2011).

Objects	m_1	m_2	m_3	m_M
m_1	Score m_{11}	Score m_{12}	Score$_{13}$	Score m_{1M}
m_2	Score m_{21}	Score m_{22}	Score m_{23}	Score m_{2M}
m_3	Score m_{31}	Score m_{32}	Score m_{33}	Score m_{3M}
m_O	Score m_{O1}	Score m_{O2}	Score m_{O3}	Score m_{OM}

Figure 6. Example of a m by m proximity matrix
Data set D = {x_1, x_2, …, x_m}, where m is the number of data points and each object is described by an m-dimensional feature vector and d_{ij} = d(x_i, x_j) with respect to some distance function d(., .).

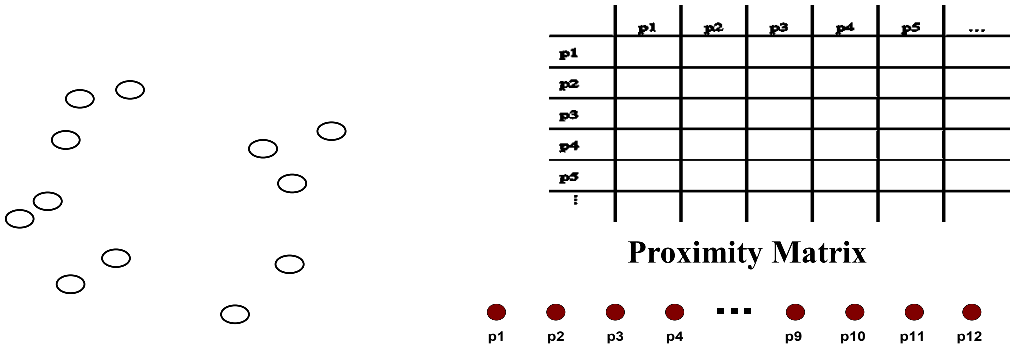

Figure 7a. Initial phase of hierarchical cluster analysis
The distances between each singleton cluster (data point) is calculated and entered into the proximity matrix.

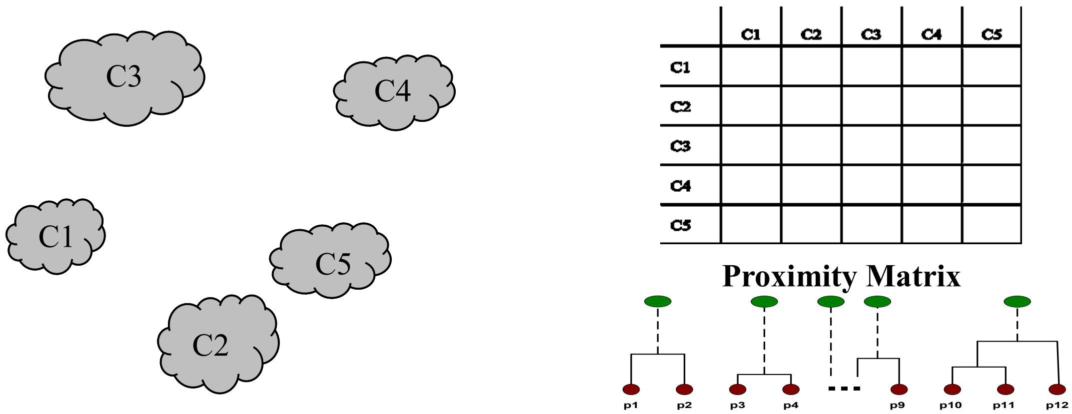

Figure 7b. Cluster formation
Singleton clusters merge and create five clusters.

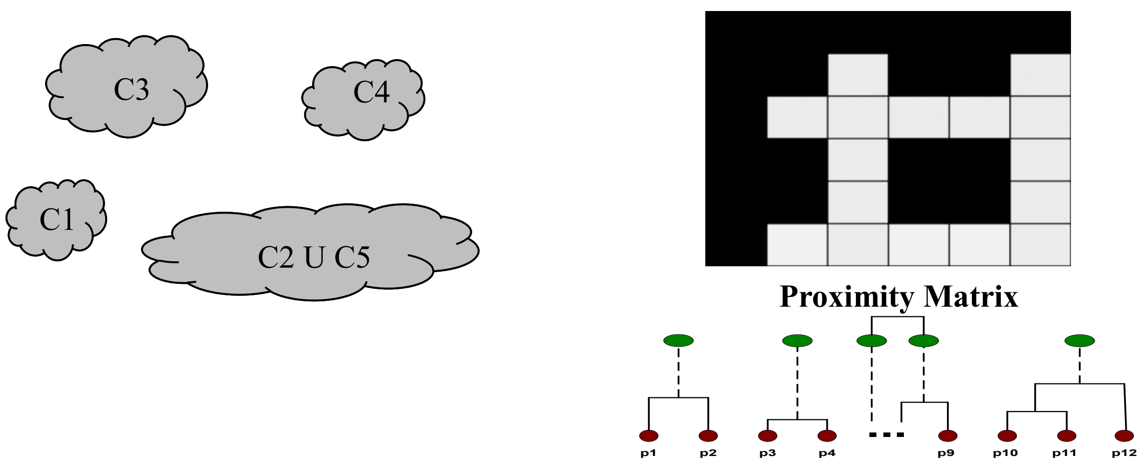

Figure 7c. Merging of clusters
The two closest clusters (2 and 5) combine and the proximity matrix is updated.

3.1.3 Cluster Linkage

All hierarchical clustering methods apply the same basic algorithm, but may differ in their measure of inter-cluster similarity, broadly termed cluster linkage. The iterative process of agglomerative cluster analysis requires the calculation of inter-cluster proximity between each cluster and the newly formed cluster at each step. There are several methods to determine the inter-cluster proximity; single linkage, complete linkage, and average linkage use a proximity matrix as input, whereas centroid linkage and Ward's method use raw data. This study uses Ward's method to determine inter-cluster similarity. Also known as the minimum sum of squares, Ward's method is unique among other forms of linkage, as it is based on an error sum of squares criterion and uses squared Euclidean distances calculated from the raw data as input (as opposed to a proximity matrix) (Everitt et al., 2011). Ward's method merges the two clusters that, when combined, minimizes the increase in the total within-cluster error sum of squares of summed over all variables (Ward Jr, 1963). The error sum of squares is given by:

$$SSE = \sum\nolimits_{i=1}^{n} (x_i - \bar{x})^2 \; ;$$

where n is the number of observations, x_i is the value of the ith observation, and $\bar{x}$ is the mean of all the observations. Ward (1963) rearranges the equation as:

$$ESS = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 .$$

3.1.3.1 Lance-Williams Recurrence Formula

Many statistical software programs use a recurrence formula to measure the proximity between clusters and expedite the clustering process for large data sets.

STATA employs the Lance and Williams (1967a) recurrence formula to calculate the distance between group *k* and newly formed (*ij*) (Lance & Williams, 1967). The Lance-Williams recurrence formula is

$$d_{k(ij)} = \alpha_i d_{ki} + \alpha_j d_{kj} + \beta d_{ij} + \gamma |d_{ki} - d_{kj}|,$$

where d_{ij} is the distance between groups i and j. Each of the cluster linkage methods have values that correspond to parameter values α_i, α_j, β, and γ, making the formula compatible to the numerous cluster linkage methods (Appendix C) (Everitt et al., 2011). Furthermore, the Lance and Williams recurrence formula uses dissimilarity measures and, if necessary, requires statistical software programs (such as STATA-13) to transform similarity measures (both binary and continuous) into dissimilarity measures prior to clustering. STATA-13 transforms similarity measures to dissimilarity measures using dissimilarity = 1 – similarity.

3.2 Partition Clustering

Partition clustering (or non-hierarchical) methods of cluster analysis begin with an initial partitioning of the data into *k* groups and progressively assigns and reassigns objects into non-overlapping clusters until objects belong to one of the *k* clusters. Partition clustering groups similar objects together in a manner that minimizes a specified error measure such as Sum of Squared Error (SSE) (Gan et al., 2007).

3.2.1 K-means Cluster Analysis

One of the most widely used partition clustering algorithms is *k*-means cluster analysis. First described by Macqueen (1967), *k*-means is a popular form of cluster analysis due to its simplicity of implementation, ability to partition large data sets, and ease in interpretation of its cluster solution and tolerance of outliers, the inclusion of irrelevant variables, and the choice of distance measure (Bailey, 1994; Gan et al., 2007).

In *k*-means cluster analysis, researchers decide upon the number of clusters a priori either at random or according to some heuristic procedure (Bailey, 1994). *K*-means cluster analysis begins with *k* seed values that act as initial cluster centers or starting centers. Objects are assigned to one of the *k* seeds whose center – or centroid – is the closest; in this study, measured by squared Euclidian distance. Once all objects are assigned to a cluster, the new cluster centroids are calculated and the iterative reassignment of objects to the nearest cluster centroid begins again. This process continues until objects no longer change clusters.

The *k*-means algorithm:

1. Input *k*, the number of clusters
2. Select *k* seed points from the data set and use them as initial centroids
3. Partition the data into *k* clusters by assigning each object to the cluster with the nearest centroid (distance – squared Euclidean between object x_i and cluster centroid).
4. Update cluster centers. Compute centroids of the clusters of the current partition. The centroid is the center (mean point) of the cluster. For each cluster *k*, the new centroid, c_k equals the mean of all points, x_i assigned to cluster S_i in previous step.
5. Repeat from step three or stop when no new assignments occur (all objects are assigned to one of the *k* clusters).

The resulting *k*-means cluster solution is one that minimizes the within-cluster error sum of squares (Hartigan & Wong, 1979).

3.2.2 Seed Values

There are options to set the initial *k* seed values including: 1. Randomly chosen objects: Where *k* objects are chosen randomly to act as starting centers for the *k* clusters; 2. The first *k*: The first *k* objects are chosen as the cluster centers for the *k* clusters; 3. The last *k* objects: The last *k* objects are chosen as the cluster centers for the *k* clusters; 4. Segments: The data is divided into *k* equal groups and the first *N*/*k* objects are assigned to the first group, the second *N*/*k* objects are assigned to the second group, until all objects are assigned to one of the *k* groups. The group means are then used as the initial starting cluster centers or seed values; and 5. Grouping variable: Where a grouping variable that defines the cluster centers based upon the means of each of the *k* groups. This study uses a grouping variable as the starting seed values.

3.2.3 The Cluster Centroid

The cluster centroid is a vector that is not necessarily a part of the original dataset and is comparable to the mean in a multi-dimensional space. In the case where an object has equal distances to two cluster centroids, the object remains in its current group (if it is the closest) or is assigned to the cluster with the lowest number of objects. The final cluster centroids reflect the mean of the typical characteristics for the members in each cluster (Bailey, 1994).

3.2.4 K-means Objective Function

The purpose of *k*-means clustering draws upon the objective function to minimize the SSE. *K*-means clustering calculates the error of each data point – according to the squared Euclidean distance of each point to the closest cluster centroid – and then calculates the total SSE. This is given by:

$$SSE = \sum_{i=1}^{K} \sum_{x \epsilon C_i} dist^2(c_i x)^2,$$

where $dist^2$ is the squared Euclidean distance between two objects in Euclidean space, x = an object; C_i = the i^{th} cluster; c_i = the centroid of cluster C_i; *c = the centroid of all points;* and K = the number of clusters.

3.3 Mixing Clustering Methodologies

Hierarchical methods and partitioning relocation methods are two forms of cluster analysis that are commonly used together to generate the most accurate cluster solution for large data sets (Tan et al., 2006). In general, hierarchical clustering methods are not ideal for extremely large data sets, but are applied to a subset of data to determine the ideal number of clusters for partition clustering and generate a grouping variable to use as the initial cluster seeds. Ward's hierarchical agglomerative cluster analysis and *k*-means cluster analysis techniques complement one another in that they both aim to minimize the SSE. Although the minimizing of square error occurs at each iteration during Ward's clustering and is a local rather than global minimum of square error objective function as it is in *k*-means clustering, Ward's clustering is often used as a robust method of initializing *k*-means clustering (Everitt et al., 2011). This study clustered a subset of data (10%) using Ward's cluster analysis, a hierarchical agglomerative algorithm, to determine the ideal number of clusters and starting seeds for a cluster analysis of the full data set using *k*-means clustering.

3.4 Validation Of The Cluster Solution

Evaluation of the cluster solution is an important component of cluster analysis. The foremost reason to validate the cluster solution is that all clustering algorithms will generate clusters from the data set even if no groups naturally exist. Demonstrating the clustering tendency of the data set will establish whether non-random structure in fact exists in the data (Tan et al., 2006). Repeated clustering of subsamples of data can demonstrate the clustering tendency of the data set and contribute to the validation of the cluster solution.

Also specific to hierarchical clustering methods is the generation of a dendrogram, which is a tree-like graphical representation of the cluster solution. The graphic portrayal of the agglomerative clustering process shows how the clusters are combined or divided at each step of the clustering procedure and can help confirm the number of groups in the cluster solution.

Reaffirmation of the optimal number of clusters is another important way to validate the cluster solution. The iterative nature of hierarchical clustering methods will cluster the data until there exists one large cluster. The employment of stopping rules helps to validate the ideal number of clusters and cluster solution by confirming when the clustering process should end (Milligan & Cooper, 1985). According to Milligan and Cooper's (1985) review of 30 methods to determine the optimal number of groups in cluster analysis, the Duda-Hart (2001) Je(2)/Je(1) index and the Calinski-Harabasz (1974) pseudo-F indices most effectively indicate the optimal number of clusters.

The Duda-Hart Je(2)/Je(1) stopping-rule index and the Calinski-Harabasz (1974) pseudo-F stopping-rule indices utilize SSE to determine cluster cohesion and cluster separation (Tan et al., 2006). Large Je(2)/Je(1) values paired with small pseudo-T-squared indicate distinct cluster formation (Duda & Hart, 1973). Large values of the pseudo-F index are an indication of distinct cluster structure, whereas smaller values indicate a lesser degree of heterogeneity between the clusters (Caliński & Harabasz, 1974).

To validate the final cluster solution researchers may also demonstrate inter-cluster heterogeneity by investigating the attributes associated with the objects in each cluster. ANOVA can determine whether the objects or attributes associated with the objects in each cluster significantly differ from those in other clusters. Variables with large F values indicate greater heterogeneity between clusters whereas those with small F values indicate lesser heterogeneity.

To the best of our knowledge, MVC research has not yet used this approach to investigate a wide range of RDBs in a large sample of drivers of all ages. Employing cluster analysis will help to determine whether patterns of RDBs exist among drivers, and if so, facilitate the identification of attributes shared by drivers who exhibit a particular pattern of driving.

Chapter 4: Manuscript 1 (Pending Submission)

Profiling Risky Driving Behaviour Among Canadian Adults – A Cluster Analysis

4.0 Abstract

Objective: To explore key factors associated with specific risky driving behaviours (RDBs) and determine whether the homogeneity hypothesis applies, whereby driving behaviours coexist among subgroups of drivers in distinct patterns.

Methods: Data was drawn from the Driving and Safety optional module of Statistics Canada's 2011 Canadian Community Health Survey, a cross-sectional nationally representative sample of 47,356 subjects aged 16 years and older who had driven a motor vehicle in the previous 12 months. The prevalence of seat belt non-compliance, cell phone-distracted driving, fatigued driving, speeding, aggressive driving, and driving under the influence of alcohol were determined. Cluster analysis was employed to generate a similarity matrix organizing the six RDBs into homologous groups based on driver involvement in each behaviour. Multiple one-way analysis of variance (ANOVA) and post-hoc pairwise comparisons using the Bonferroni correction evaluated between group differences to confirm whether distinct groups existed within the data.

Results: Cluster analysis revealed five heterogeneous groups of drivers based on behaviour patterns, which suggested that three patterns of RDBs exist among drivers. The five-cluster solution included two very risky subgroups, the Poly-risk Drivers and Egocentric Drivers, one moderately risky subgroup, the Average Drivers, one unique subgroup that refrained from all RDBs aside from seat belt non-compliance, the Beltless Drivers, and a subgroup of drivers who refrained from all RDBs, the Cautious Drivers. Profiles were generated based on the characteristics associated with the drivers in each subgroup and indicated that in addition to age and sex, lifestyle, health, and cultural factors contributed to RDBs.

Conclusion: Exploration of the factors associated with six RDBs through cluster analysis demonstrated that drivers who engaged in a specific RDB were not distinct from other risky drivers, but belonged to a larger profile of drivers who engaged in multiple RDBs. As has been established, young to middle age and male sex were determined to be important factors associated with RDBs. Lifestyle, health, and cultural factors were also identified as important considerations for future research of drivers who engage in RDBs.

4.1 Road Safety in Canada

Motor vehicle collision-related morbidity and mortality are major population health concerns. The health and economic burdens associated with motor vehicle collisions (MVCs) have prompted much traffic safety research to explore the factors and risks associated with drivers involved in MVCs (Transport Canada, 2013). Traffic safety research has identified human factors as the most influential of all factors associated with MVCs, more so than engineering factors (vehicle design and safety features) and environmental factors (road design and crash-protective roadside objects). More than any other human factor, risky driving behaviours (RDBs) contribute to the majority of MVCs (Evans, 1991). Efforts to improve traffic safety, however, have traditionally focussed on vehicle and roadway design and road safety education rather than changes in driver behaviour (Gielen & Sleet, 2003). Countermeasures successful in maintaining behaviour changes and improving population health have established a clear understanding of the determinants of behaviour and have focussed on lifestyle changes (Green & Kreuter, 1999; McGinnis, Williams-Russo, & Knickman, 2002). Therefore, identification of the factors associated with RDBs is vital to identifying high risk drivers and improving traffic safety.

There is considerable research dedicated to RDBs, MVCs, and road safety; yet, research on the factors shared by drivers who engage in RDBs is minimal and has often focused on the contributions of driving behaviours to crash-risk. Much of this work lacks methodological rigor, varies in RDB terminology, and is limited to subpopulations of drivers not generalizable to the larger driving population. The majority of research on driving behaviour has focused on seat belt non-compliance, speeding, and driving under the influence of alcohol (DUIA), and more recently, cell phone-distracted driving and aggressive driving; there is a paucity of research on fatigued driving (Caird, Willness, Steel, & Sciafa, 2008; Di Milia, et al., 2011; Smart & Mann, 2002).

Although factors such as young age and male sex are known to contribute to engagement in RDBs, there is a lack of consensus among researchers regarding many of the other factors associated with RDBs (Evans, 1991; Transport Canada, 2014). Furthermore, given that risk-taking behaviours in general rarely occur in isolation and health-compromising behaviours that contribute to injury often co-occur, it is plausible that multiple RDBs co-occur among drivers (Anderson & Mellor, 2008; Dohmen et al., 2011; Jessor, 1987; McDonald, Sommers, & Fargo, 2014a; Petridou et al., 1997). To the best of our knowledge, no research has explored whether the "homogeneity hypothesis" holds true in the context of RDBs. It remains unclear whether drivers who engage in a specific RDB are distinct from other risky drivers or if they belong to a larger profile of drivers who engage in multiple RDBs. A comprehensive understanding of the factors associated with a broad range of RDBs, as well as identifying whether individual factors are associated with specific RDBs or if they are common to multiple interrelated driving behaviours, is needed to better profile high-risk drivers. Such research will clarify the determinants of RDBs and allow traffic safety efforts to better target interventions to particular subpopulations of risky drivers.

Thus far, research exploring the factors associated with RDBs has identified sociodemographic characteristics such as age, sex, socioeconomic status (SES), marital status, geographic location (urban or rural), and race as correlates of risky driving (Golias & Karlaftis, 2001; Sahai, Pitblado, Bota, & Rowe, 1998; Transport Canada, 2014). The

Canadian Council of Motor Transportation Administrators (CCMTA) survey of 3,888 Canadian drivers identified male sex, young age and higher income as factors associated with cell phone-distracted driving, fatigued driving, and aggressive driving and associations between male sex and low-income and DUIA and seat belt non-compliance. Results also demonstrated an association between female sex and speeding, contradicting the majority of research indicating that males are more likely to speed than females (Vanlaar, Simpson, Mayhew, & Robertson, 2008; Vingilis, 2007).

Although research has demonstrated associations between particular sociodemographic factors and RDBs, much of the research thus far has included a limited number of driver characteristics. Despite clear associations between poor mental health and engagement in many other risk-taking behaviours such as smoking, unsafe sexual behaviours, poor diet, and physical inactivity, minimal research has explored the association between health factors and RDBs (Murphy et al., 2014; Scott & Happell, 2011; Ziedonis et al., 2008). Some evidence of an association between psychiatric distress, stress, depression, smoking, and alcohol consumption and RDBs have been observed, particularly with respect to aggressive driving and DUIA. However, such research examined crash risk or injury and less is known about the association between measures of mental health and driving behaviour (Linden et al., 2010; Mann et al., 2010; Smart, Asbridge, Mann, & Aldaf, 2003; G. Stoduto et al., 2008; Wickens, Smart, & Mann, 2014; Wickens et al., 2013).

Furthermore, research has demonstrated that health-compromising behaviours that contribute to injury often co-occur within individuals (Anderson & Mellor, 2008; Dohmen et al., 2011; Jessor, 1987; McDonald et al., 2014a; Petridou et al., 1997).

However, research on the factors associated with RDBs has rarely included factors reflecting on whether risky drivers also take risks in other areas of their lives. Thus far, research has been limited to young drivers or has investigated risk perception, personality characteristics, attitudes towards RDBs, or sensation seeking rather than specific factors representing risk propensity such as riding with a drinking driver (RWDD), binge drinking, or the number of injuries incurred in the previous year (Dahlen, Martin, Ragan, & Kuhlman, 2005; Lajunen & Parker, 2001; Ulleberg & Rundmo, 2003).

According to Jessor (1986), young drivers who engage in one form of RDB are likely to engage in other forms of risky driving or patterns of RDBs. Research on whether drivers' exhibit patterns of driving behaviours have focussed primarily upon the impact of developmental and psychosocial factors among youth and adolescents (Jessor, 1987; Jonah & Dawson, 1987). It remains unclear whether RDBs are interrelated and coexist among risky drivers as patterns of RDBs.

Previous research has focussed upon a specific RDB as the dependent variable, and did not compare contributing factors across a broad range of driving behaviours. For example, Sahai, et al. (1998) used self-reported survey data and showed that while controlling for a number of demographic variables, drivers who engaged in speeding (OR = 2.04, 95% CI: 1.71 – 2.44) and DUIA (OR = 2.43, 95% CI: 2.23 – 2.60) were over twice as likely to be seat belt non-compliant than non-speeders and drivers who did not DUIA, respectively. Other research using self-reported driving behaviour and observed highway driving data by Zhao, Reimer, Mehler, D'Ambrosio, & Coughlin (2013) showed that drivers who reported frequent cell phone use while driving were also likely to have engaged in speeding and aggressive driving. Zhao, et al. (2013) used ANOVA analysis of cell phone use on a number of driving performance measures and found that compared to rare cell phone users, frequent users had a higher mean velocity [$F(1, 102) = 381.6$, $p = 0.00$], spent more time in the left lane [$F(1, 100) = 11.7$, $p = 0.001$), and engaged in a number of measures indicative of aggressive driving such as more rapid throttle accelerations [$F(1, 102) =5.5$, $p = 0.02$], hard braking events [$F(1, 102) = 4.6$, $p = 0.03$], and sudden non-directional accelerations [$F(1, 102) = 12.6$, $p = 0.001$] independent of driver age and sex. Although such examples are valuable explorations of factors associated with specific RDBs, they restrict the comparison of a broad range of factors associated with multiple RDBs.

Research employing cluster analysis to investigate patterns of RDBs is minimal. Studies that have used this approach have reported cluster solutions that contained small numbers of distinct groups of drivers, but were, for the most part, based upon the clustering of personality measures or restricted to young drivers (Deery & Fildes, 1999; Lucidi et al., 2010; Ulleberg, 2001; Vassallo et al., 2007). Vassallo, et al. (2007) employed a cluster analysis of the frequency of speeding, seat belt non-compliance, fatigue, DUIA, and driving under the influence of an illegal drug to investigate the longitudinal precursors of RDBs among young adult drivers (19 and 20 years). The cluster solution identified three distinct groups of low, moderate, and high-risk drivers with the high-risk subgroup of drivers more likely to have engaged in speeding, seat belt non-compliance, and DUIA than drivers in the moderate-risk and low-risk clusters. Longitudinal measures of behaviour problems, social skills, and peer relationships uniquely profiled each of the three clusters of drivers, with higher values of these measures associated with the high-risk subgroup of drivers (Vassallo et al., 2007). Such findings suggest patterns of RDB, but similar research inclusive of drivers of all ages is required to determine whether such patterns of RDBs exist in the larger driving population.

The objectives of this study were twofold. First, using population data, to examine the factors associated with six RDBs – seat belt non-compliance, cell phone-distracted driving, fatigued driving, speeding, aggressive driving, and DUIA – and use cluster analysis to determine whether subgroups of drivers exist. Second, if subgroups of RDBs were present, to determine whether the drivers in each subgroup shared common characteristics. If so, sociodemographic, health-related, and other risk-taking behaviour variables would help to characterize the subgroups of drivers and validate the cluster solution by further differentiating the subgroups of drivers in the cluster solution. To the best of our knowledge, this was a novel approach to explore the factors associated with RDBs and to examine whether drivers who engaged in a specific RDB were distinct from other risky drivers or if they belong to a larger profile of drivers who engaged in multiple RDBs.

4.2 Methodology

Data for this study was drawn from a subsample of the 2011 Canadian Community Health Survey (CCHS). The CCHS is a cross-sectional nationally representative sample of 131,486 individuals aged 12 and older living in 139,841 private dwellings covering 117 health regions in all Canadian provinces and territories. Residents of Indian Reserves, Canadian Forces Bases, institutional accommodations and some individuals living in extremely remote areas were excluded from the sampling frame. Information pertaining to health status, health care utilization and determinants of health were collected from the Canadian population to generate estimates at the health region level. In addition to the CCHS core component, all health regions were offered participation in 50 additional optional modules, giving provinces the option of selecting content addressing the health priorities of their health region or province. The subsample of data used in this study was drawn from the CCHS optional module Driving and Safety, which contained questions related to a broad range of risky driving behaviours (Statistics Canada, 2011).

4.2.1 Data Source

Statistics Canada collected data between January 2009 and December 2010. CCHS coverage was in the range of 98% in the provinces, 97% in the Northwest Territories, 90% in the Yukon and 71% in Nunavut. Data was collected via face-to-face and telephone interviews and achieved an overall response rate of 72.3%. The provinces of Newfoundland, Ontario, Alberta and the Yukon participated in the Driving and Safety optional module of the CCHS. The sample was comprised of 59,163 subjects; 3,768 from Newfoundland, 42,495 from Ontario, 11,618 from Alberta and 1,282 from the Yukon Territory. For the purposes of this study, subjects of interest included those 16 years and older who had driven a motor vehicle (including a car, truck, or van) in the previous 12 months. The total subsample of data for this study contained 47,356 subjects.

4.2.2 Outcome Variables

Outcome variables included six questions related to RDBs from the Driving and Safety optional module of the CCHS, comprised of five ordinal variables: 1. Seat belt non-compliance; 2. Cell phone-distracted driving; 3. Fatigued driving; 4. Speeding; and 5. Aggressive driving; and one binary variable, DUIA. Measurement was scaled according to the response categories for each outcome variable as displayed in the CCHS, with lower measures indicating increased frequency, magnitude, or engagement and higher measures indicating reduced frequency, magnitude, or non-engagement in each risky driving behaviour: seat belt use, the regularity of seat belt use (*never*, *rarely*, *sometimes*, *always*, coded 1 – 4); distracted driving, how often the respondent used a cell phone while driving (*often*, *sometimes*, *rarely*, or *never*, coded 1 - 4); fatigued driving, driving while feeling tired (*often*, *sometimes*, *rarely*, or *never,* coded 1 - 4); speeding (*much faster*, *a little faster*, *about the same speed*, or *a little slower*, or *much slower*, coded 1 - 5); aggression (*much more aggressively*, a *little more aggressively*, *about the same*, *a little less aggressively*, or *much less aggressively*, coded 1 - 5); and DUIA, defined as having two or more drinks within one hour prior to driving a motor vehicle (*yes* or *no*, coded 0, 1). Research has demonstrated that hands-free cell phone distracted driving reduces vehicle control and may not be a safe alternative to hand-held devices (McCartt, Hellinga, & Bratiman, 2006). Due to variation in driver attitudes, perceptions of risk, and motivations to use hands-free devices as a safer alternative to hand-held devices, hands-free cell phone-distracted driving was not measured in the present study (White, Eiser, & Harris, 2004; Zhou, Wu, Rau, &Zhang, 2009). A description of the primary outcome variables and response frequencies is in Appendix D.

4.2.3 Independent Variables

Sociodemographic, health-related, and other risk-taking behaviours were examined across response categories of the RDBs. Missing categories were included for variables with more than four percent missing data, as missing rates of five percent or less have been shown to be inconsequential to statistical inferences (Schafer, 1999). Sociodemographic variables included *age* (<=20 years, 21-25, 26-40, 41-55, 56-70, and >=71, coded 1-6), *sex* (female or male, coded 0,1), *geography* (rural or urban, coded 0,1), *race* (White, non-White, or missing, coded 1-3), *immigrant status* (non-immigrant or immigrant, coded 0,1), *marital status* (married or common law or not married, coded 0,1), *education* (< secondary education, secondary education, some post-secondary education, post-secondary graduate, or missing, coded 1-5), *household income* ($0-$19,999, $20,000-$39,999, $40,000-$59,999, $80,000 or more, or missing, coded 1-6), and *employment status* (employed or unemployed, coded 0,1).

Health-related variables included *self-perceived physical health* (healthy or unhealthy, coded 0,1), *self-perceived mental health* (positive or negative, coded 0,1), *stress* (lower levels of stress or higher levels of stress, coded 0,1), *diagnosis of a mood disorder* (no mood or mood, coded 0,1), *diagnosis of an anxiety disorder* (no anxiety or anxiety, coded 0,1), and *satisfaction with life* (ten-point scale 0 = very dissatisfied, 10 = very satisfied, coded 1-10).

Variables associated with other risk-taking behaviours unrelated to driving included: *smoking* (non-smoker or smoker, coded 0,1), *alcohol* (no alcohol or alcohol, coded 0,1), *binge drinking* (non-drinker, drinker, non-binger, and binger, coded 1-3), *riding with a drinking driver* (*RWDD*) (no RWDD or RWDD, coded 0,1), and *number of injuries* (not injured, one time, or two or more, coded 1-3) (Appendix E).

4.3 Analysis

All analyses were performed using STATA 12.0 and SPSS 2012 (*IBM SPSS statistics for MacIntosh,* 2012; StataCorp., 2011). Probability sampling weights were used to produce population estimates at the health region level for Canadian drivers 16 years and older. The weight variable was rescaled so that the average weight was equal to one to account for the unequal probability of selection. This was achieved by dividing the original weight by the average of the original weights (245.94) for the number of respondents. In addition, STATA's robust option was employed to adjust standard errors for survey design effects resulting from the CCHS complex sampling design. All estimates ascertained from the master CCHS files were based on a minimum of 10 observations.

Data analysis occurred in three stages. First, descriptive statistics evaluated the prevalence of the six RDBs and the distribution of the independent variables across the response categories for each of the RDBs to determine whether associations existed between driver characteristics and engagement in risky driving. Differences between groups of categorical variables were evaluated according to chi-square values and t-tests compared sample means for continuous measures.

Second, cluster analysis identified groups with varying profiles of risky driving. Cluster analysis organized the data into homologous groups (or clusters) based respondents' common responses to the six questions pertaining to risky driving (Aldenderfer, 1984; Sokal & Sneath, 1963). Ward's cluster analysis was performed on a subset (10%) of data (randomly generated by STATA-12) to determine the appropriate number of groups for subsequent *k*-means analysis (Everitt, Landau, Leese, & Stahl, 2011). Ward's method assessed cluster membership based on the total sum of squared deviations from the mean of a cluster and merged the two clusters that produced the smallest increase in the error sum of squares according to squared Euclidian distance (Everitt, et al., 2011). Given the absence of an objective measure to determine the optimal cluster solution, the number of groups in the Ward's cluster solution was determined by assessing the changes in the agglomerative coefficient according to the Duda-Hart (2001) Je(2)/Je(1) index and the Calinski-Harabasz (1974) pseudo-F indices and through examination of its visual representation, the resulting tree dendrogram (Everitt, et al., 2011). The within-cluster means from the Ward's cluster solution were then saved and used as starting centroids in the subsequent *k*-means cluster analysis (Caliński & Harabasz, 1974; Duda & Hart, 1973; Duda, Hart, & Stork, 2012; Milligan & Cooper, 1985).

K-means cluster analysis, an iterative partitioning method, was then employed to organize the full sample (n = 47,356) into the predetermined number of clusters generated by Ward's cluster analysis using the within-cluster means as starting centroids. The two clustering methods are complimentary, as Ward's cluster analysis is useful for determining the optimal number of clusters, but performs best with small sample sizes, and *k*-means cluster analysis is useful for confirming the cluster composition using large samples, but the number of starting centroids (clusters) must be known a priori (Everitt et al., 2011).

Third, the cluster solution was further validated according to the differences in the means associated with each RDB and descriptive characteristics across clusters. Validation of a cluster solution is achieved when a cluster solution contains high degrees of inter-cluster heterogeneity (differences in the means of each variable across clusters) and intra-cluster homogeneity (similarity of variables within each cluster) (Everitt et al., 2011). A series of one-way analyses of variance (ANOVA) assessed whether significant differences in the means existed across the subgroups of drivers. F-test scores evaluated the contribution of the individual variables; the Bonferroni method was employed for post-hoc pairwise comparison to reduce the possibility of Type 1 errors and significant group differences were noted at the $p = 0.05$ level.

4.4 Results

Descriptive characteristics of the subsample were comparable to the CCHS data set in its entirety and representative of the Canadian population (Appendix F) (Statistics Canada, 2011). Results showed a large percentage of missing data for income (35.3%). Smaller, yet noteworthy percentages of missing data were found for education (7.9%) and race (5.6%).

The distribution of driver engagement across the categories of each RDB demonstrated a range of involvement in risky driving (Table 1). The prevalence of each RDB was calculated according to any involvement in the specified behaviour (Figure 8). Results showed that the prevalence of seat belt non-compliance (5.3%) and DUIA (9.6%) were quite low compared to cell phone-distracted driving (40.1%), speeding (30.9%), aggressive driving (30.9%), and the most prevalent RDB, fatigued driving (66.5%).

A total of 21 driver characteristics were distributed across the subcategories of each RDB. Chi-square tests revealed significant differences ($p < 0.001$) between the categories of descriptive characteristics and involvement in each of the six RDBs. Similarly, t-test scores showed significant differences ($p < 0.001$) between the means of continuous descriptive characteristics and involvement in each of the RDBs (Tables 2 – 7). Associations between driver characteristics and engagement in each of the RDBs demonstrated that many factors of were common to multiple RDBs.

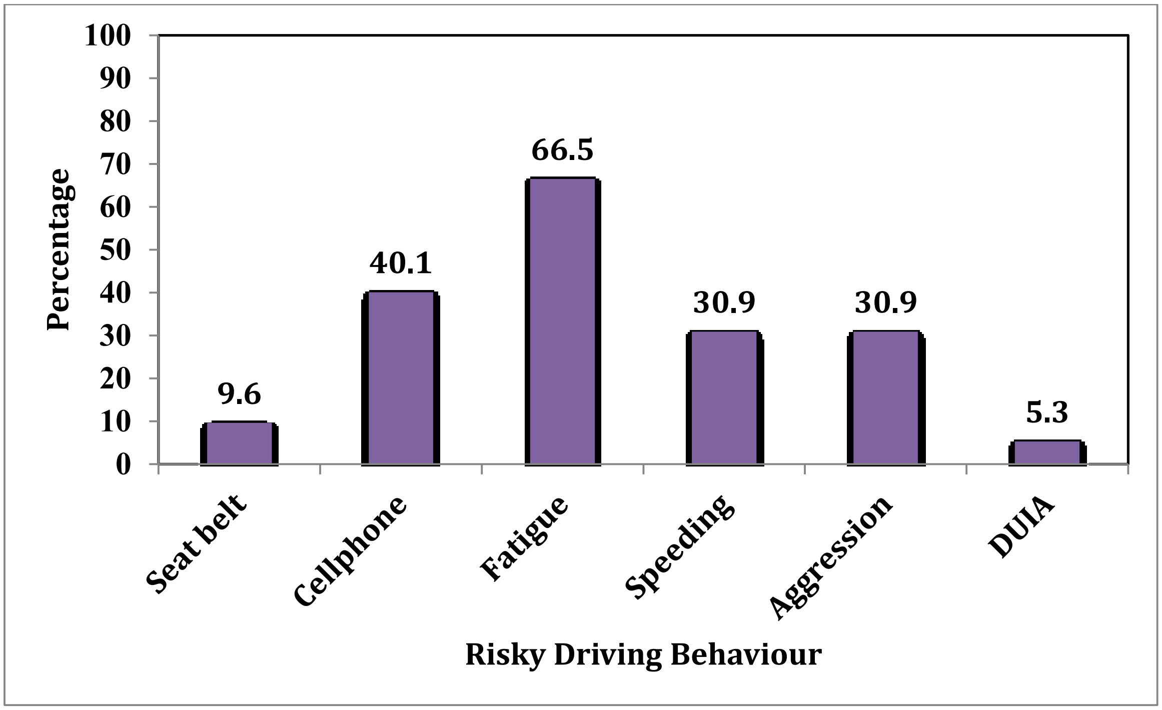

Figure 8. Prevalence of risky driving behaviours (Canadian Community Health Survey, 2011)
Prevalence was calculated according to any involvement in each RDB. Measurement of the six RDBs included: 1. Seat belt non-compliance: *always*, *most of the time*, *rarely*, or *never*; 2. Cell phone-distracted driving: *often*, *sometimes*, *rarely*, or *never*; 3. Fatigued driving: *often*, *sometimes*, *rarely*, or *never*; 4. Speeding: *much faster*, *a little faster*, *about the same*, *a little slower*, or *much slower*; 5. Aggressive driving: *much more aggressively*, *a little more aggressively*, *about the same*, *a little less aggressively*, or *much less aggressively*; and 6. DUIA: *yes* or *no*.

4.4.1 Age and Sex

Results showed higher proportions of drivers aged 26-40 reported more frequent (*often*) fatigued driving and more extreme speeding (*much faster*) and aggressive driving (*much more aggressively*) than drivers in other age categories. For example, the percentage of drivers aged 26-40 years (12.1%) who reported *often* driving while fatigued was higher than drivers under 20 years of age (5.2%), 21- 25 years (12.0%), 41-55 years (8.9%), 56-70 years (3.6%), and 71 years and older (0.8%). Chi-square tests demonstrated significant differences between the categories of *age* and level of engagement in fatigued driving [$\chi 2(15, N = 47{,}356) = 4{,}391, p < .001$] (Table 4). Results for cell phone-distracted driving and DUIA showed the age category 21-25 years contained the highest proportions of drivers who reported *always* for cell phone-distracted driving and *yes* to DUIA, while a higher proportion of drivers 71 years and older reported *never* using a seat belt than drivers in all other age categories. Results for driver *sex* showed higher proportions of males than females reported seat belt non-compliance (*never*), frequent phone-distracted driving (*often*) and fatigued driving (*often*), and more extreme speeding (*much faster*), and aggressive driving (*much more*), and engagement in DUIA (*yes*).

4.4.2 Sociodemographic Factors

Results demonstrated significant associations between the sociodemographic factors and each of the RDBs. Compared to other levels of educational attainment, higher proportions of drivers with some post-secondary education reported fatigued driving (*often*), speeding (*much faster*), aggressive driving (*much more*), and DUIA (*yes*). A higher proportion of drivers with secondary education reported *often* engaging in cell phone-distracted driving and drivers with less than secondary education reported *never* for seat belt compared to other education categories. Results for annual household income showed that compared to drivers who reported lower incomes, drivers with incomes over $80,000 contained the highest proportion of drivers who reported *often* for cell phone-distracted driving and fatigued driving, *much faster* for speeding, *much more* for aggressive driving, and *yes* for DUIA. An association between low income and seat belt non-compliance was also noted; compared to all other levels of annual household income categories, drivers who earned less than $20,000 contained the highest proportion of drivers who reported *never* using a seat belt. Results for employment demonstrated higher proportions of employed drivers than unemployed drivers *often* engaged in cell phone-distracted driving and fatigued driving, drove *much faster* and *much more aggressively* and engaged in DUIA. In contrast, a significantly higher proportion of unemployed drivers than employed drivers reported *never* using a seat belt.

Results for the health-related variables showed the mean values for satisfaction with life were lower among drivers who reported *much faster* (M = 7.8, SE = 1.8) compared to drivers who reported *much slower* (M = 9.1, SE = 1.6) (Table 5) in addition to drivers who reported driving *much more* aggressively (M = 7.8, SE = 1.8) compared to drivers who reported much less aggressively (M = 9.0, SE = 1.7) (Table 6). Results demonstrated associations between higher levels of stress, diagnosis of a mood disorder, and diagnosis of an anxiety disorder and multiple forms of risky driving. Greater proportions of drivers with higher levels of stress reported *often* engaging in cell phone-distracted driving, *often* driving while fatigued driving, driving *much faster,* and *much more* aggressively, compared to drivers with *lower levels of stress*. Higher proportions of drivers with a mood disorder reported *never* using a seat belt, *often* driving while fatigued, driving *much faster*, and *much more* aggressively than other drivers compared to drivers without a mood disorder. Similar results were found for anxiety. Results also showed higher proportions of drivers who reported their health as *unhealthy* and their mental health as *negative* reported *never* using a seat belt than those who reported *healthy* and *positive*, respectively. A slightly higher proportion of drivers with negative mental health also reported driving *much more* aggressively than other drivers, however, a much higher proportion of drivers with negative mental health reported driving *much less* aggressively compared to drivers with positive mental health.

Results also suggested associations between other risk-taking behaviours and RDBs. Higher proportions smokers than non-smokers reported *often* for cell phone-distracted driving, *often* engaging in fatigued driving, *much faster* for speeding, *much more* for aggressive driving, and engagement in DUIA. Compared to drivers who do not drink alcohol, higher proportions of drinkers reported *never* using a seat belt, *often* for cell phone-distracted driving, driving *much faster* than other drivers and DUIA. Higher proportions of drivers who binge drink, engage in RWDD, and incurred two or more injuries in the previous 12 months reported *often* for cell phone-distracted driving, *much faster* for speeding, *much more* for aggression, *often* for fatigued driving, and *yes* for DUIA compared to drivers who who do not binge drink, do not engage in RWDD, and have not been injured, respectively. Finally, a higher proportion of drivers who did not engage in RWDD also reported seat belt non-compliance (*never*) compared to those who reported engagement in RWDD.

4.4.3 The Cluster Solution

Ward's cluster analysis revealed five homologous clusters in the data. The cluster stopping rules verified five distinct clusters. The large Je(2)/Je(1) value paired with small pseudo-T-squared value from the Duda-Hart Je(2)/Je(1) stopping rule and the large value of the pseudo-F from the Calinski/Harabasz stopping rule provided evidence of five distinct clusters within the data (Appendix G). The dendrogram generated post-clustering provided a visual representation of the hierarchical clustering process and further confirmed the five-cluster solution (Figure 9). In addition, repeated Ward's hierarchical cluster analysis was performed on multiple 10% subsamples, the majority of which generated 5-cluster solutions, providing further support for the five-cluster solution presented here.

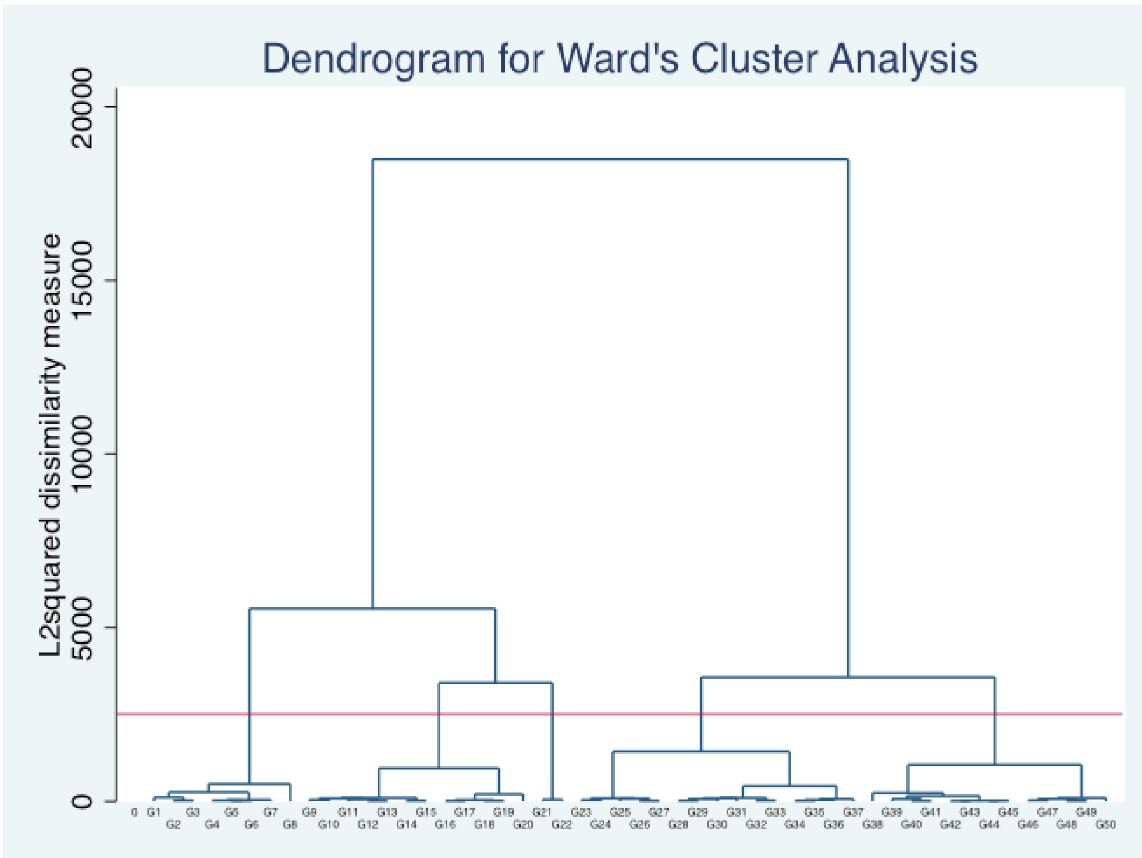

Figure 9. Dendrogram generated by Ward's cluster analysis.
The resulting dendrogram provides a visual representation of the hierarchical cluster process, with the nodes representing the clusters and the heights of the stems representing the squared Euclidean distances where clusters merge (Everitt et al., 2011). The final 50 mergers between clusters are displayed and clearly show five distinct clusters (red line).

The *K*-means procedure classified the data into five distinctly heterogeneous clusters, each with a high level of homogeneity among the individuals within each cluster. The mean value for each of the RDBs was calculated by cluster. Low mean values indicated more frequent (*always* or *often*), more extreme or greater intensity (*much faster* or *much more aggressively*) or engagement in (*yes*) the risky driving behaviour. A series of one-way analysis of variance (ANOVA) evaluated between-group differences (Table 8). Results indicated significant differences between the means of each RDB ($p<.001$) across the five clusters – seat belt use [$F(4, 47{,}351) = 46{,}288.0$, $p < .000$]; cellphone-distracted driving [$F(4, 47{,}351) = 19{,}807.2$, $p < .000$]; fatigued driving [$F(4, 47{,}351) = 78{,}490.4$, $p < .000$]; speeding [$F(4, 47{,}351) = 85{,}333.3$, $p < .000$]; aggressive driving [$F(4, 47{,}351) = 81{,}991.4$, $p < .000$]; and DUIA [$F(4, 47{,}351) = 328.4$, $p < .000$]. Post-hoc comparisons using the Bonferroni test confirmed that all pairwise comparisons for each of the RDBs were significantly different between clusters with two exceptions. The mean values for cell phone-distracted driving in cluster 1 ($M = 3.66$, $SD = 0.47$) and cluster 2 ($M = 3.66$, $SD = 0.72$) did not significantly differ from one another ($p = 1.0$), and DUIA in cluster 2 ($M = 0.02$, $SD = 0.14$) and cluster 3 ($M = 0.01$, $SD = 0.01$) did not significantly differ from one another ($p = 0.27$).

The clusters identified were as follows (names assigned by author):

1. The Average Drivers (Cluster 1, $n = 14{,}179$) – this cluster contained drivers who generally refrained from RDBs; drivers reported comparable speed and aggression to other drivers, reported seat belt compliance, and no DUIA, but reported minimal engagement in cell phone-distracted driving and fatigued-driving.
2. The Cautious Drivers (Cluster 2, n=15,791) – was comprised of drivers who reported little or no engagement in RDBs; drivers reported minimal cell phone-

distracted driving, seat belt compliance, never engaged in fatigued driving or DUIA, and drove slower and less aggressively than other drivers.

3. The Beltless Drivers (Cluster 3, n = 2,083) – drivers in this cluster refrained from all RDBs aside from seat belt non-compliance.
4. The Egocentric Drivers (Cluster 4, n = 5,558) – the second riskiest cluster contained drivers who often used a cell phone while driving and engaged in DUIA, and reported moderate levels of speeding and aggressive driving.
5. The Poly-risk Drivers (Cluster 5, n = 9,745) – the riskiest subgroup of drivers engaged in most RDBs; drivers in this cluster often drove while fatigued, much faster, much more aggressively than other drivers, occasionally engaged in DUIA, but reported minimal cell phone-distracted driving.

The distribution of sociodemographic, (Table 9), health-related (Table 10), and other risk-taking behaviours (Table 11) were examined to note patterns or trends that further differentiated the five subgroups of drivers. A second set of ANOVA tests demonstrated significant differences between the means for all descriptive characteristics associated with the drivers in each cluster ($p < .001$). Post-hoc pairwise comparisons using the Bonferroni tests revealed significant differences in the means of most descriptive variables. Significant differences between all five clusters were found for age [$F(4, 47351) = 544.6$, $p < .000$], income [$F(4, 31{,}000) =317.2$, $p < .000$ $p < .000$], employment [$F(4, 47351) =1620.5$, $p < .000$], self-perceived health status [F(4, 47,351) =316.8, $p < .000$], binge drinking [F(4, 47351) = 1,266.0, $p < .000$], and RWDD [F(4, 47351) =341.7, $p < .000$] (Tables 12-14).

Profiles were created based on the characteristics associated with the drivers in each of the five clusters. Most driver characteristics associated with the Average Drivers were comparable to the sample population, although the Average Drivers had a marginally higher mean age ($M = 49.5$, $SE = 0.2$) than the sample population ($M = 44.8$, $SE = 0.01$) and a higher proportion of immigrants (33.4% vs. 22.1%), drivers with post-secondary education (77.3% vs. 73.5%) (Table 12), drivers with positive mental health (96.0% vs. 93.1%), and drivers who consume alcohol (84.5 vs. 80.0%) (Table 13).

The Cautious Drivers had a mean age of 54.8 years ($SE = 0.2$) and approximately 40% of drivers in this cluster were over the age of 56 years. Compared to other clusters, the Cautious Drivers had the highest proportion of female (52%) and non-White (22%) drivers. Aside from the Beltless Drivers, the Cautious Drivers reported lower educational attainment than other clusters, with over seven percent of the Cautious Drivers reporting less than secondary school education, while the percentage of drivers with less than secondary education in each of the other subgroups of drivers was five percent or lower. Similarly, over four percent of drivers in this subgroup reported an annual household income of less than $20,000, compared to three percent or lower in other clusters. The proportion of Cautious Drivers that reported positive mental health (95%) was slightly higher than that of the sample population (93%). In addition, these drivers had the lowest proportion of all clusters of drivers who reported higher levels of stress (58%) and two or more injuries in the previous 12 months (2.8%) (Table 14).

The Beltless Drivers had a mean age of 55.1 years ($SE = 0.5$) with 47% of cluster members aged 56 years and older. Relative to other clusters, the Beltless Drivers contained the highest proportion of married drivers (57.2%), employed drivers (31.6%), and drivers who reported post-secondary degrees (39.2%) and incomes of $80,000 or more (12.2%). However, this cluster contained considerably higher proportions of missing data for the variables education (45.0%) and income (66.1%) compared to the other clusters. The health-related characteristics associated with the Beltless Drivers were particularly noteworthy. Compared to all other clusters, the Beltless Drivers contained the highest proportion of drivers who reported their self-perceived health as unhealthy (31.3%), negative mental health (54.9%), a mood disorder (10.5%), and an anxiety disorder (7.5%). Yet, this subgroup had the highest life satisfaction of all clusters (M = 9.0, SD = 0.0) in addition to the lowest proportion of smokers (15%) and drivers who consume alcohol (32.5%). Finally, results indicated that the Beltless Drivers avoided other risk-taking behaviours. The proportion of Beltless Drivers who reported binge drinking (9.0%) and RWDD (1.7%) was significantly lower than other clusters, in which at least 27% of drivers reported binge drinking and at least 6% reported RWDD.

Cluster 4, the Egocentric Drivers, had the lowest mean age 37.3 years (SE = 0.2) of all clusters and compared to all other clusters, contained the highest proportion of drivers who were male (63%), living in urban areas (85%), who reported annual household incomes over $80,000 (44%), and who were employed (93%). A high percentage of these drivers also reported completing post-secondary school education (78%), were unmarried (38%) and White (82%). Results for health-related variables showed that although the majority of the Egocentric Drivers reported being healthy (94.2%), 23% of drivers in this subgroup were smokers and almost 90% of drivers consumed alcohol. In addition, compared to the other subgroups of drivers, the Egocentric Drivers had the second highest proportion of drivers who reported higher levels of stress (75%) and the highest proportion of drivers who reported binge drinking (60%), RWDD (20%), and two or more injuries in the past 12 months (5%).

Finally, the distribution of descriptive characteristics uniquely profiled the Poly-risk Drivers (Cluster 5). The Poly-risk Drivers had a mean age of 44.2 years (SE = 0.2), which was the second lowest mean age of all clusters after the Egocentric Drivers. Compared to the four other clusters, the Poly-risk Drivers contained the highest proportion of drivers who reported post-secondary education (79%) and, after the Egocentric Drivers, the second highest proportion of drivers of drivers who reported an annual household income of $80,000 or more (39%). Results for the health-related factors showed that of all clusters, the Poly-risk Drivers contained the highest proportion of drivers who reported higher levels of stress (77%) and aside from the Beltless Drivers, drivers who reported a mood disorder (8.0%), and drivers who reported an anxiety disorder (5.2%). Furthermore, the Poly-risk Drivers had the lowest life satisfaction (M = 7.8, SD = 0.2) of all clusters. Compared to other clusters, the Poly-risk Drivers contained the second highest proportion (after the Egocentric Drivers) of binge drinkers (46.4%), drivers who engaged in RWDD (11.7%), and drivers who reported two or more injuries in the past 12 months (4.5%).

In general, the sociodemographic, health related, and risk-taking behaviours associated with each cluster further differentiated the five subgroups of drivers and provided external validity to the cluster solution. In addition, repeated clustering of subsamples confirmed the five-cluster solution. Overall, the five-cluster solution was a significant improvement over the null model that assumed no clusters were present in the data; the five-cluster results indicated that it was a good model fit for the data.

4.5 Discussion

The focus of this research was to examine the factors associated with a broad range of RDBs in a large population of drivers of all ages. Cluster analysis was employed to determine whether patterns of RDBs existed and, if so, the study tested whether specific factors are common to these multiple interrelated RDBs. Based upon the current MVC literature, it was anticipated that the riskiest drivers were more likely to be young and male. It was hypothesized that cluster analysis would reveal distinct patterns of RDBs, particularly patterns that included speeding, aggression, cell phone-distracted driving, and DUIA, and that the demographic profiles of each subgroup of drivers would further differentiate the clusters.

The preliminary analysis of this study suggested that although a relatively small proportion of the driving population engaged in frequent or extreme degrees of RDBs, a much larger proportion of drivers admitted to at least some engagement. The range of involvement in fatigued driving exemplified this trend; eight percent of drivers reported *often* driving while fatigued, while almost 60% reported *sometimes* or *rarely*. The prevalence of drivers who engaged in RDBs to at least some degree emphasized the importance of clarifying the determinants of RDBs so that traffic safety efforts may more effectively identify risky drivers and implement measures to reduce their prevalence and improve traffic safety (Gielen & Sleet, 2003).

4.5.1 Factors Associated With RDBs

Exploration of the characteristics associated with risky drivers found that many factors were common to multiple driving behaviours. Consistent with previous research, the present study demonstrated associations between sex and engagement in RDBs with higher proportions of males than females engaging in RDBs, particularly cell phone-distracted driving and DUIA (Asbridge, Brubacher, & Chan, 2013; Fernandes, Hatfield, & Job, 2010; Jonah, 1990; Transport Research International Documentation (TRID), 2013; Vingilis & Wilk, 2010). This study also demonstrated associations between driver age and engagement in RDBs with higher proportions of middle-aged drivers reporting cell phone-distracted driving, fatigued driving, aggressive driving, and speeding compared to drivers in older and young age categories. Although much research has demonstrated associations between younger age and many RDBs, the findings of the present study complement more recent literature demonstrating that middle aged drivers also engage in most forms of risky driving (Asbridge et al., 2013; Di Milia et al., 2011; Transport Canada, 2005; Transport Research International Documentation (TRID), 2013).

Another interesting finding was that drivers over the age of 71 contained the highest proportion of seat belt non-compliant drivers of all age categories. Such findings contradict the majority of previous research identifying young age as a factor in seat belt non-compliance (Canadian Council of Motor Transport Administrators, 2014; Sahai et al., 1998; Vingilis & Wilk, 2010; Wilson, 1990). The high proportion of older seat belt non-compliant drivers suggested a cohort effect, perhaps resulting from the average age of licensure of older drivers, which was likely prior to seat belt safety campaigns.

In addition to age and sex, a number of other sociodemographic factors were associated with factors related to health and other risk-taking behaviours and multiple RDBs. Consistent with previous research, the present study demonstrated links between higher levels of educational attainment, income, and employment and most RDBs

(Canadian Council of Motor Transport Administrators, 2014). In contrast, this study indicated an association between low socioeconomic status and seat belt non-compliance. The high proportions of low income, low educational attainment, and unemployed drivers, in addition to the high mean age of these drivers suggests that retirement status may have affected study results, particularly those for seat belt non-compliance. In addition, although surveys have indicated that non-response to income-related questions is between 20 and 40%, the highest proportions of missing data for income in this study were found among drivers who abstained from RDBs (Juster & Smith, 1997). Given that the missing data for income may not have been random, future research may consider employing a technique for missing data or using education as a proxy for income when exploring associations between socioeconomic status and RDBs. Regardless, the findings for employment and education reaffirmed the associations suggested between socioeconomic status and engagement in RDBs in the present study.

A number of health factors were also common to multiple RDBs. This study showed that higher levels of stress, diagnosis of a mood disorder, diagnosis of an anxiety disorder, smoking, and lower life satisfaction were each linked to fatigued driving, speeding, and aggressive driving. Associations between high levels of stress and cell phone-distracted driving, as well as associations between smoking and cell phone-distracted driving were also apparent. Previous research has demonstrated that psychiatric disorders may impede driver performance and may lead to an increase in the risk of MVC (Mann et al., 2010; Wickens et al., 2014; Wickens et al., 2013). The GHQ measure of psychiatric distress, a measure of nonpsychotic psychiatric illnesses such as anxiety, psychological distress, and social functioning, has been linked to driver aggression (Smart et al., 2003). The preliminary results of this study did not show a clear association between negative mental health and aggressive driving. Although a slightly higher proportion of drivers with negative mental health reported driving *much more* aggressively than drivers with positive mental health, a substantially higher proportion of drivers with negative mental health reported driving *much less aggressively* than drivers with positive mental health. These findings may indicate that specific measures of mood or anxiety disorders more accurately reflect psychiatric distress compared to the more subjective and broad variable capturing negative mental health in this study. Regardless, the findings of the present study highlighted the importance of further research on the contribution of mental health factors to RDBs.

According to the literature, risk-taking behaviours rarely occur in isolation and health-compromising behaviours that contribute to injury often co-occur within individuals (Anderson & Mellor, 2008; Dohmen et al., 2011; Jessor, 1987; McDonald et al., 2014a; Petridou et al., 1997). The findings of this study also suggested associations between a general propensity towards taking risks and engagement in RDBs, particularly binge drinking, RWDD, and previous injuries (risks) and the engagement in cell phone-distracted driving, speeding, aggressive driving, and DUIA (RDBs). These findings are reflective of the concept of problem behaviour syndrome in the context of risky driving, whereby drivers who engage in one RDB are likely to engage in others, perhaps due to, in part, a lower perception of risk compared to safer drivers (Jessor, 1987; Jonah & Dawson, 1987). The present study demonstrated that this concept may be applied to drivers of all ages and not limited to young drivers.

4.5.2 The Five-Cluster Solution

Recognizing that multiple RDBs shared common factors, the main analysis in this study utilized cluster analysis to determine whether natural patterns of RDBs existed among Canadian drivers. The five distinct subgroups of drivers in the cluster solution indicated that patterns of driving behaviours were present confirming the homogeneity hypothesis in the context of risky driving. The Poly-Risk Drivers and Egocentric Drivers highlighted two very risky patterns of driving behaviours. The Poly-Risk Drivers (20.6%) – arguably the riskiest subgroup of drivers who engaged in a pattern of speeding, aggressive driving, and fatigued driving – may represent drivers who may be more deviant in general. The second subgroup of risky drivers, the Egocentric Drivers (11.7%), demonstrated a second pattern of RDBs indicative of drivers who prioritize their needs in the moment over safety. These drivers frequently engaged in cell phone-distracted driving and DUIA more than any other cluster, and engaged in speeding and aggressive driving. This subgroup of drivers complements previous research linking narcissism and engagement in impulsive and self-defeating behaviours such as DUIA and aggressive driving in addition to a grandiose sense of self and a lowered perception of risk (Lustman, Wiesenthal, & Flett, 2010; Miller, et al., 2009). The group of Average Drivers (30.0%) revealed a third, moderately risky pattern of behavior, not as dangerous as the Poly-risk Drivers or Egocentric Drivers, but not as safe as the Cautious Drivers. The Average Drivers engaged in fatigued driving and cell phone-distracted driving, yet reported driving at speeds and levels of aggression comparable to other drivers. This pattern of RDBs suggested that although some drivers may not have reported extreme degrees of RDBs, they still engaged in behaviours that endangered themselves, passengers, and other road users. Such findings highlight the need for further research and education on the dangers of RDBs, particularly cell phone-distracted driving and fatigued driving.

In contrast to these clusters of risk-taking drivers, the Cautious Drivers (33.3%) engaged in minimal RDBs and reported driving more slowly and less aggressively than other drivers, while the smallest cluster, the Beltless Drivers (4.4%) exhibited no RDBs other than seat belt non-compliance. Although the cluster of Beltless Drivers was small, it significantly differed from the four other subgroups in the cluster solution and efforts to improve seat belt compliance may benefit from targeting this distinct subset of drivers. In addition, these findings also suggested that the associations between the independent variables and seat belt compliance (in the preliminary analysis of this study) may have somewhat accurately differentiated the seat belt non-compliant drivers from the rest of the driving population, despite the aforementioned limitations of such associations.

Aside from the small subgroup of Beltless Drivers, the clustering of RDBs indicated that the majority of drivers (62.3%) reported engagement – to varying degrees – in most RDBs (the Poly-risk Drivers, Egocentric Drivers, and Average Drivers), while a smaller percentage of drivers (33.3%) refrain from risky driving (the Cautious Drivers). Such findings support previous research indicating that the majority of drivers engage in multiple risky driving behaviours, and not just a single RDB (Jessor, 1987; Jonah & Dawson, 1987; Zhao, et, al., 2013). Furthermore, the three clusters of risky drivers in the present study were moderately similar to the three-cluster solution found by Lucidi et al., (2010) of risky, worried, and careful drivers and suggested a distinction between assertive risky drivers (the Poly-risk Drivers and Egocentric Drivers) and drivers who did not intentionally, or irregularly, engaged in RDBs (the Average Drivers).

Previous research has noted patterns of seat belt non-compliance, speeding, and DUIA, in addition to patterns of cell phone-distracted driving and aggressive driving (Sahai et al., 1998; Zhao et al., 2013). Although the patterns of multiple RDBs revealed by cluster analysis in this study did not include seat belt non-compliance, the broad range of RDBs and population of drivers unrestricted by age provided a more comprehensive investigation of patterns of RDBs generalizable to the larger driving population.

4.5.3 Cluster Profiles

Another valuable finding of this research was the distribution of descriptive characteristics across the five-cluster solution. The unique characteristics associated with each of the five clusters further contributed to the heterogeneity between clusters in the five-cluster solution. Moreover, the cluster profiles verified the homogeneity hypothesis in the context of risky driving, whereby drivers who engaged in multiple RDBs belonged to a larger profile of drivers with similar characteristics. Age and sex, in addition to factors related to lifestyle, health, and culture uniquely profiled each of the subgroups of risky drivers.

4.5.3.1 Age and Sex

The high proportions of young male drivers in the two high-risk clusters, the Poly-risk Drivers and the Egocentric Drivers, compared to drivers in other clusters support previous research demonstrating associations between young age and male sex and engagement in risky driving (Canadian Council of Motor Transport Administrators, 2014; Jonah, 1986). The profile of the Egocentric Drivers, contrasted with that of the Cautious Drivers, demonstrated this finding. The Egocentric Drivers had the lowest mean age and highest proportion of males of all five clusters, while the Cautious Drivers contained the highest proportion of females of all five clusters, in addition to the second highest mean value for age (second only to the Beltless Drivers).

The high mean age associated with the Beltless Drivers cluster contradicted previous studies that showed seat belt compliance increases with age (Sahai et al., 1998). Although the prevalence of seat belt non-compliance was lower than that of other RDBs, the subgroup of Beltless Drivers in the cluster solution suggested that the benefits of seat belt use were not well understood by this subgroup of older drivers (Canadian Council of Motor Transport Administrators, 2014). The Beltless Drivers' seat belt non-compliance may be explained by their year of licensure – prior to the strict enforcement of seat belt laws and knowledge of benefits of seat belt safety – or that risk perception or attitudes towards risk may play a role in seat belt non-compliance (McCartt, Mayhew, Braitman, Ferguson, & Simpson, 2009; Simpson & Mayhew, 1992). The lack of a variable to represent year of licensure, may have limited study results.

4.5.3.2 Lifestyle

According to the literature, lifestyle plays an important role in young drivers' involvement in risky driving (Jessor, Turbin, & Costa, 1997; Jessor, 1987). The findings of the present study indicated that lifestyle may influence engagement in RDB of drivers of all ages. First, the present study demonstrated an association between affluence and engagement in RDBs, which compliments previous research linking high socioeconomic status to risky driving (Canadian Council of Motor Transport Administrators, 2014; Smart et al., 2003). The two riskiest clusters of drivers, the Poly-risk Drivers and Egocentric Drivers, both contained high percentages of drivers with high incomes, levels of education, employed drivers, and drivers with higher levels of stress compared to other clusters. This trend of affluence, employment, and higher levels of stress among the two riskiest subgroups of drivers signified that a busy working lifestyle may play a role in risky driving. In contrast, the two safest clusters of drivers, the Cautious Drivers and Beltless Drivers, were characterized by unemployment, low educational attainment and household incomes, and minimal stress. These findings may be explained by research demonstrating that drivers no longer in the workforce engage in fewer RDBs (Bhatti et al., 2008; Charlton, Oxley, Fildes, Oxley, & Newstead, 2003; Charlton et al., 2006). Research of over 10,000 drivers by Bhatti, et al. (2008) showed that compared to drivers still in the workforce, retired drivers were 66% more likely to have discontinued fatigued driving (OR= 2.12, p = <0.001) and cell phone-distracted driving (OR = 1.74, p = 0.006). Given these findings, future research may benefit from exploring the contribution of retirement status on the engagement in RDBs.

Second, compared to the Poly-risk Drivers, the Egocentric Drivers were slightly younger and more likely to be unmarried, employed, have high levels of education and household incomes, and live in urban locations. This driver profile may suggest that the Egocentric Drivers lead fast-paced lifestyles common among young adults who are not yet settled into family life. Alternatively, given that the Egocentric Drivers had a high mean value for stress and contained a high proportion of smokers, these drivers may also have reflected a lifestyle common among young urban professionals with demanding jobs. It was also speculated that the high household incomes and educational attainment of the drivers in this cluster may have been influenced by shifts in Canadian workforce demographics, although verification of such a phenomenon was beyond the scope of this study (Burke & Ng, 2006). Shifts in workforce demographics due to the aging workforce and technological advances are nonetheless an important consideration for future MVC research.

Third, it was apparent that the propensity to take risks among the riskiest subgroups of drivers was not limited to driving behaviours, but was also manifested in other areas of their lives. In contrast to the safer Cautious Drivers, the Poly-risk Drivers and Egocentric Drivers engaged in multiple RDBs and reported binge drinking, RWDD, and two or more injuries in the past 12 months. This risk-taking, problem behaviour lifestyle supported previous research that demonstrated that risky drivers were also likely to have exhibited risk-taking behaviours in other areas of their lives (Jessor, 1987a; Musselwhite, 2006).

4.5.3.3 Health Factors

The findings of this study also suggested associations between health factors and RDBs. For example, the Poly-risk Drivers had the lowest life satisfaction and the highest proportion of drivers of all clusters who reported a diagnosis of a mood disorder and higher levels of stress. Higher levels of stress and aggressive driving also characterized the cluster of Egocentric Drivers, albeit to a lesser degree than the Poly-risk Drivers. Such results were consistent with previous literature that showed psychiatric distress and stress as two factors common among aggressive drivers, as well as consistent with traits associated with mental health disorders, RDBs, and crash risk (Mann, Smart, Stoduto, Adlaf, & Ialomiteanu, 2004; Matthews et al., 1998; Smart et al., 2003; Williams, Tregear, & Amana, 2011).

The profile of the cluster of Beltless Drivers also indicated an association between health factors and risky driving. These seat belt non-compliant drivers were significantly more likely to report their health status as 'unhealthy' and mental health status as 'negative' than drivers in other clusters. In addition, similar to the Poly-risk Drivers, large percentages of Beltless Drivers reported diagnoses of either a mood disorder or anxiety disorder compared to other clusters. This may be attributed to the high mean age of these drivers, as many health-related issues become apparent and are diagnosed with age (Kessler, Olfson, & Berglund, 1998; Kessler et al., 2005; Olfson, Kessler, Berglund, & Lin, 1998; Wang et al., 2007). Also of note was the high life satisfaction among the Beltless Drivers. Such findings, however, are not surprising, as much research has demonstrated that satisfaction with life remains constant among middle aged and older individuals up until the age of 70 years, in spite of health problems (Baird, Lucas, & Donnellan, 2010).

The inclusion of other health-related variables in this research such as illicit drug use and problem gambling may have further confirmed the association between health factors and RDBs. However, the provincial variation in participation for the optional modules of the CCHS limited this study to variables common to the four provinces. In addition, factors such as sex, current health status, and differences in cultural judgements of questions related to health may have affected participants' responses and data validity (Idler & Benyamini, 1997).

4.5.3.4 Sociodemographic Factors

The findings of this study also suggest associations between cultural factors and RDBs. Aside from immigrant drivers who reported *sometimes* engaging in cell phone-distracted driving, a lower proportion of participants who identified as a non-White compared to White or as an immigrant to Canada compared to a non-immigrant reported engagement in RDBs. The cultural makeup of the Cautious Drivers exemplified this trend, with a large proportion of immigrant drivers and drivers who identified as visible minorities belonging to this subgroup of safe drivers.

Although cultural factors may have played a role in driving behaviours, reporting bias may have affected the responses, whereby non-White and immigrant drivers may have underreported RDBs more so than White and non-immigrant drivers. Factors such as social desirability and fear of reprisal on participants' responses may have limited the data used in this study, as these factors are known to influence participants' responses to sensitive questions related to illegal behaviours such as impaired driving (Lajunen & Summala, 2003).

4.6 Considerations for Further Analyses

First, this study did not standardize variables prior to the clustering procedures. Although there is debate as to whether to standardize variables prior to cluster analysis, standardization is known to be particularly beneficial when variables have widely differing scales or large standard deviations (Bible, Datta, & Datta, 2013; Miligan & Cooper, 1998). This study did not standardize variables due to the small range of the variable scales and to avoid masking natural patterns of RDBs in the data. Future research may benefit from comparing the cluster results from both unstandardized and standardized data. Second, the characteristics associated with the five-cluster solution were limited to CCHS variables common to all four provinces included in this study. Variables known to impact RDBs such as attitudes towards risk were not included in the profiling of the five-cluster solution and may limit study results. The inclusion of such a measure may benefit future research. Finally, this study's use of self-reported data to reflect RDBs may not be as reliable as studies based on objective MVC-related morbidity or mortality data. Despite this potential limitation, research shows self-reported driving behaviour is a valid measure of driving behaviour (Lajunen & Summala, 2003; West, French, Kemp, & Elander, 1993).

The cluster analysis of RDBs using population-based data proved to be a viable and unique method to explore the RDBs. The present study examined associations between a broad range of driver characteristics and a comprehensive range of RDBs in a large population of drivers unrestricted by age. The profiles associated with each of the five subgroups of drivers revealed in this study indicated that in addition to young to middle age and male sex, factors related to lifestyle, health, culture, and other risk-taking behaviours were associated with specific patterns of RDBs. These findings warrant further analyses of the factors associated with RDBs and research of patterns of RDBs. Viewing crash risk as a function of RDBs highlighted the importance of identifying the human factors associated with RDBs, particularly factors associated with lifestyle and others amenable to the prevention of RDBs. This approach to MVC research may benefit programs such as the GDL program by tailoring their education according to program location or driver demographics.

Table 1
Frequency (%), mean (*sd*), and prevalence of risky driving behaviours among Canadian adult drivers

Risky driving behaviour	Response category	Frequency (%)	M^a (SD)^b	Prevalence^c (%)
Seat belt use (1 – 4)	Always	42,837.6 (90.5)	3.8 (0.7)	4.4 - 9.6
	Most of the time	2,013.0 (4.3)		
	Rarely	409.9 (0.9)		
	Never	2,095.5 (4.4)		
Cellphone use (1 – 4)	Often	3,369.4 (7.1)	3.3 (1.0)	7.1 - 40.1
	Sometimes	5,773.7 (12.2)		
	Rarely	9,831.1 (20.8)		
	Never	28,381.7 (59.9)		
Fatigued driving (1 – 4)	Often	3,803.8 (8.0)	3.0 (0.9)	8.0 - 66.5
	Sometimes	10,864.7 (22.9)		
	Rarely	16,841.8 (35.6)		
	Never	15,845.8 (33.5)		
Speeding (1 – 4)	Much faster	3,803.8 (8.0)	3.0 (1.0)	8.0 - 30.9
	A little faster	10,864.7 (22.9)		
	The same	16,841.8 (35.6)		
	A little slower	13,920.3 (29.4)		
	Much slower	1,925.5 (4.1)		
Aggressive driving (1 – 5)	Much more aggressively	3,796.2 (8.0)	3.0 (1.0)	8.0 - 30.9
	A little more aggressively	10,838.5 (22.9)		
	The same	16,795.9 (35.5)		
	A little less aggressively	13,784.4 (29.1)		
	Much less aggressively	2,141.1 (4.5)		
DUIA (1 – 2)	Yes	2,496.4 (5.3)	0.95 (0.2)	5.3
	No	44,859.6 (94.7)		
Total		47,356 (100)		

^{a} Mean, ^{b} Standard deviation, ^{c} Prevalence of each RDB or a range in prevalence where the lower limit reflects the most frequent (*never*, *always*, *often*) or highest degree (*much faster*, *much more aggressively*) of engagement in RDBs and the upper limit reflects inclusion of reports of all degrees of engagement (*most of the time*, *sometimes*, *rarely*, *a little faster*, *a little more aggressively*).

Table 2
Percent distribution of driver characteristics across seat belt compliance ($p < .05$)

Descriptive characteristic		*n* %	Always	Most of the time	Rarely	Never	Total	*df*	χ2
Age	<=20	7.4	87.6	4.5	0.8	7.0	100	15	1,406
	21-25	8.5	89.3	6.9	0.8	3.0	100		
	26-40	26.2	90.8	4.8	1.2	3.2	100		
	41-55	30.5	92.3	4.2	0.9	2.6	100		
	56-70	19.4	92.0	3.2	0.5	4.2	100		
	71+	8.0	82.5	2.1	0.4	15.0	100		
Sex	Female	47.5	93.3	2.0	0.4	4.3	100	3	671
	Male	52.5	87.9	6.3	1.3	4.5	100		
Geographic area	Rural	16.6	87.4	7.2	1.3	4.2	100	3	216
	Urban	83.4	91.1	3.7	0.8	4.5	100		
Immigrant	No	74.5	89.4	4.9	1.0	4.6	100	3	217
	Yes	25.5	93.5	2.2	0.4	3.8	100		
Race	White	77.8	91.6	4.7	1.0	2.7	100	6	4,718
	Non-White	16.5	94.1	2.0	0.3	3.6	100		
	Missing	5.6	63.8	4.9	0.9	30.5	100		
Marital Status	Married or common-law	65.8	91.4	3.9	0.9	3.8	100	3	112
	Not married	34.3	88.6	4.9	0.9	5.6	100		
Education	Less than secondary	4.4	88.0	5.0	1.3	5.7	100	12	4,034
	Secondary graduate	9.5	90.9	5.0	1.3	3.4	100		
	Some post-secondary	4.7	91.1	6.0	1.1	2.4	100		
	Post-secondary graduate	73.5	92.7	4.4	0.8	2.4	100		
	Missing	7.9	70.0	6.4	0.8	24.7	100		
Income	$0-$19,999	2.9	90.7	3.1	1.0	5.2	100	15	1,012
	$20,000-$39,999	7.9	91.2	3.3	0.8	4.7	100		
	$40,000-$59,999	9.9	91.6	4.7	1.0	2.7	100		
	$60,000-$79,999	10.5	93.2	4.5	0.4	1.8	100		
	$80,000 or more	33.6	92.7	4.5	1.2	1.7	100		
	Missing	35.3	87.1	4.1	0.7	8.2	100		
Employment	Employed	76.3	92.2	4.9	1.0	2.0	100	3	2,282
	Unemployed	23.7	84.9	2.3	0.5	12.3	100		

Descriptive characteristic		Seat belt							
		n %	Always	Most of the time	Rarely	Never	Total	*df*	χ2
Self-perceived health	Healthy	89.5	91.5	4.3	0.9	3.4	100	3	937
	Unhealthy	10.5	82.1	4.1	1.0	12.9	100		
Mental health	Positive	93.1	92.7	4.3	0.9	2.2	100	3	7,691
	Negative	6.9	60.1	4.2	0.9	34.9	100		
Stress	Lower levels of stress	32.6	91.0	3.5	0.7	4.8	100	3	47
	Higher levels of stress	67.4	90.2	4.6	1.0	4.2	100		
Mood	No mood	93.5	90.6	4.3	0.9	4.3	100	3	51
	Mood	6.5	88.0	4.0	1.1	6.9	100		
Anxiety	No anxiety	95.4	90.6	4.2	0.9	4.3	100	3	50
	Anxiety	4.6	87.5	4.5	0.6	7.4	100		
Smoking	Non-smoker	79.6	91.1	3.5	0.7	4.7	100	3	355
	Smoker	20.4	87.9	7.0	1.7	3.4	100		
Alcohol	No alcohol	20.0	82.2	2.6	0.7	14.6	100	3	2,942
	Alcohol	80.0	92.5	4.7	0.9	1.9	100		
Binge drinking	Non-drinker	20.4	82.2	2.6	0.7	14.6	100	6	3,464
	Drinker, non-binger	40.1	94.4	2.7	0.5	2.5	100		
	Binger	39.5	90.8	6.7	1.4	1.2	100		
RWDD	No	89.9	90.8	3.6	0.8	4.8	100	3	585
	Yes	10.1	87.4	9.9	1.7	1.0	100		
Injuries	Not injured	85.4	90.7	3.9	0.8	4.6	100	6	184
	1 injury	10.9	90.0	5.4	1.6	3.1	100		
	2 or more injuries	3.7	86.7	9.0	0.7	3.6	100		
Continuous measures									*t*
Age2	*M*[a]	44.8	44.7	40.8	41.3	51.5		47,355	571
	SE[b]	0.1	16.7	15.7	14.6	23.3			
Satisfaction with life	*M*	8.1	8.0	7.8	7.6	9.0		47,355	1,087
	SE	0.0	1.6	1.6	2.2	1.6			

[a] Mean, [b] Standard error of the mean

Table 3
Distribution of driver characteristics across cellphone-distracted driving ($p < .05$)

Descriptive characteristic		Cell phone-distracted driving							
		N %	Often	Some-times	Rarely	Never	Total	*df*	χ2
Age	<=20	7.4	6.2	12.5	18.5	62.7	100	15	4,647
	21-25	8.5	13.7	20.5	23.3	42.5	100		
	26-40	26.2	10.4	17.2	26.7	45.7	100		
	41-55	30.5	7.0	12.5	22.7	57.8	100		
	56-70	19.4	3.1	5.9	16.2	74.8	100		
	71+	8.0	0.3	0.8	4.5	94.4	100		
Sex	Female	47.5	4.7	10.6	20.2	64.5	100	3	601
	Male	52.5	9.3	13.6	21.3	55.8	100		
Geographic area	Rural	16.6	6.3	10.5	20.3	62.9	100	3	44
	Urban	83.4	7.3	12.5	20.9	59.3	100		
Immigrant	No	74.5	7.7	12.6	21.8	57.9	100	3	255
	Yes	25.5	5.3	10.9	17.9	65.9	100		
Race	White	77.8	7.5	12.3	21.6	58.6	100	6	222
	Non-White	16.5	5.5	12.9	19.1	62.5	100		
	Missing	5.6	6.2	8.5	14.4	70.9	100		
Marital Status	Married or common-law	65.7	6.8	11.6	21.1	60.4	100	3	41
	Not married	34.3	7.7	13.2	20.1	59.0	100		
Education	Less than secondary	4.4	5.2	4.4	9.0	81.4	100	12	717
	Secondary graduate	9.5	7.7	11.1	17.8	63.4	100		
	Some post-secondary	4.7	7.3	14.2	19.1	59.4	100		
	Post-secondary graduate	73.5	7.2	13.0	22.5	57.2	100		
	Missing	7.9	6.1	8.8	15.6	69.5	100		
Income	$0-$19,999	2.9	4.3	7.5	11.9	76.3	100	15	1,413
	$20,000-$39,999	7.9	3.8	8.5	13.1	74.6	100		
	$40,000-$59,999	9.9	5.6	10.1	18.0	66.3	100		
	$60,000-$79,999	10.4	5.9	13.8	22.6	57.6	100		
	$80,000 or more	33.5	9.8	14.8	25.6	49.8	100		
	Missing	35.3	6.3	11.0	18.8	63.8	100		

Descriptive characteristic		Cell phone-distracted driving							
		N %	Often	Some-times	Rarely	Never	Total	*df*	χ2
Employment	Employed	76.3	8.9	14.5	23.9	52.6	100	3	3,454
	Unemployed	23.7	1.3	4.7	10.6	83.4	100		
Self-perceived health	Healthy	89.5	7.5	12.7	21.5	58.3	100	3	467
	Unhealthy	10.5	3.6	7.8	14.6	74.0	100		
Mental health	Positive	93.1	7.4	12.6	21.3	58.7	100	3	418
	Negative	6.9	3.2	6.5	13.6	76.7	100		
Stress	Lower levels of stress	32.6	4.8	10.2	18.6	66.4	100	3	465
	Higher levels of stress	67.4	8.2	13.2	21.8	56.8	100		
Mood	No mood	93.5	7.2	12.4	20.7	59.7	100	3	40
	Mood	6.5	5.4	9.7	21.0	63.9	100		
Anxiety	No anxiety	95.4	7.1	12.3	20.8	59.8	100	3	14
	Anxiety	4.6	7.2	10.1	19.6	63.2	100		
Smoking	Non-smoker	79.6	6.7	11.9	20.8	60.6	100	3	62
	Smoker	20.4	8.5	13.4	20.8	57.3	100		
Alcohol	No alcohol	20.0	3.5	8.0	14.8	73.8	100	3	972
	Alcohol	80.0	8.0	13.2	22.2	56.5	100		
Binge drinking	Non-drinker	20.4	3.5	8.2	14.9	73.4	100	6	2,715
	Drinker, non-binger	40.1	4.7	9.8	19.2	66.3	100		
	Binger	39.5	11.4	16.7	25.4	46.5	100		
RWDD	No	89.9	6.1	11.4	20.2	62.4	100	3	1,348
	Yes	10.1	16.3	19.6	25.8	38.3	100		
Injuries	Not injured	85.4	6.8	12.0	20.3	60.9	100	6	160
	1 injury	10.9	8.0	12.8	23.3	56.0	100		
	2 or more injuries	3.7	11.6	14.5	24.8	49.1	100		
Continuous measures									*t*
Age2	*M*[a]	44.8	37.0	37.8	41.0	48.5	47,355		571
	SE[b]	0.1	12.5	13.0	14.0	18.2			
Satisfaction with life	*M*	8.1	8.2	8.0	8.1	8.1	47,355		1,087
	SE	0.0	1.5	1.5	1.4	1.7			

[a] Mean, [b] Standard error of the mean.

Table 4
Percent distribution of driver characteristics across fatigued driving ($p < .05$)

Descriptive characteristic		Fatigue							
		N %	Often	Some-times	Rarely	Never	Total	*df*	$\chi 2$
Age	<=20	7.4	5.2	16.8	32.8	45.2	100	15	4,391
	21-25	8.5	12.0	28.7	40.0	19.4	100		
	26-40	26.2	12.1	28.2	36.5	23.2	100		
	41-55	30.5	8.9	26.4	36.7	28.0	100		
	56-70	19.4	3.6	16.7	36.1	43.6	100		
	71+	8.0	0.8	7.2	25.0	67.1	100		
Sex	Female	47.5	7.7	21.1	34.9	36.3	100	3	180
	Male	52.5	8.3	24.6	36.2	30.9	100		
Geographic area	Rural	16.6	8.6	24.2	35.5	31.7	100	3	18
	Urban	83.4	7.9	22.7	35.6	33.8	100		
Immigrant	No	74.5	8.7	23.2	37.6	30.5	100	3	602
	Yes	25.5	6.2	22.1	29.6	42.2	100		
Race	White	8.4	23.5	37.9	30.2	8.4	100	6	999
	Non-White	6.2	23.1	28.9	41.8	6.2	100		
	Missing	8.5	15.3	22.5	53.6	8.5	100		
Marital Status	Married or common-law	65.8	7.9	23.1	36.0	33.1	100	3	11
	Not married	34.3	8.3	22.7	34.8	34.2	100		
Education	Less than secondary	4.4	4.7	13.5	28.9	52.9	100	12	1,287
	Secondary graduate	9.5	5.8	21.8	32.9	39.6	100		
	Some post-secondary	4.7	8.7	21.2	35.5	34.7	100		
	Post-secondary graduate	73.5	8.6	24.4	37.6	29.5	100		
	Missing	7.9	7.1	17.3	23.8	51.8	100		
Income	\$0-\$19,999	2.9	5.8	18.0	27.8	48.3	100	15	1,508
	\$20,000-\$39,999	7.9	5.8	18.6	30.0	45.6	100		
	\$40,000-\$59,999	9.9	7.0	22.0	35.1	35.9	100		
	\$60,000-\$79,999	10.5	8.3	25.8	35.9	30.1	100		
	\$80,000 or more	33.6	10.7	26.9	38.9	23.5	100		
	Missing	35.3	6.4	19.9	34.3	39.3	100		
	Employed	76.3	10.0	26.7	37.5	25.9	100		

Descriptive characteristic		Fatigue							
		N %	Often	Some-times	Rarely	Never	Total	*df*	χ2
Employment	Unemployed	23.7	1.9	11.0	29.2	58.0	100	3	4,416
Self-perceived health	Healthy	89.5	8.0	23.5	36.6	31.9	100	3	486
	Unhealthy	10.5	7.9	18.3	26.6	47.1	100		
Mental health	Positive	93.1	7.8	23.5	36.8	31.9	100	3	849
	Negative	6.9	10.6	15.6	19.1	54.7	100		
Stress	Lower levels of stress	32.6	4.3	17.0	36.3	42.3	100	3	1,276
	Higher levels of stress	67.4	9.8	25.8	35.2	29.2	100		
Mood	No mood	93.5	7.8	22.9	35.8	33.5	100	3	58
	Mood	6.5	11.4	23.6	31.9	33.2	100		
Anxiety	No anxiety	95.4	7.9	22.9	35.8	33.4	100	3	42
	Anxiety	4.6	11.3	23.3	31.4	34.1	100		
Smoking	Non-smoker	79.6	7.3	22.8	35.8	34.2	100	3	176
	Smoker	20.4	11.1	23.7	34.7	30.6	100		
Alcohol	No alcohol	20.0	5.4	16.2	26.0	52.4	100	3	1,900
	Alcohol	80.0	8.7	24.6	37.9	28.7	100		
Binge drinking	Non-drinker	20.4	5.4	16.2	26.2	52.2	100	6	3,028
	Drinker, non-binger	40.1	6.5	21.0	36.3	36.2	100		
	Binger	39.5	10.9	28.4	39.6	21.1	100		
RWDD	No	89.9	7.4	21.9	35.4	35.3	100	3	801
	Yes	10.1	13.4	31.9	37.5	17.2	100		
Injuries	Not injured	85.4	7.5	22.6	35.1	34.8	100	6	304
	1 injury	10.9	10.9	23.6	39.7	25.7	100		
	2 or more injuries	3.7	11.9	27.9	35.2	25.0	100		
Continuous variables									*t*
Age2	*M*[a]	44.8	38.6	41.4	44.0	49.6	47,355		571
	SE[b]	0.1	12.5	14.1	16.2	19.5			
Satisfaction with life	*M*	8.1	7.8	7.9	8.1	8.2	47,355		1,087
	SE	1.6	1.8	1.5	1.5	1.8			

[a] Mean, [b] Standard error of the mean.

Table 5
Percent distribution of driver characteristics across speeding (p < .05)

Descriptive characteristic		Speeding							
		N (%)	Much faster	Little faster	The same	Little slower	Much slower	*df*	χ2
Age	<=20	7.4	5.2	16.8	32.8	38.7	6.6	15	4,984
	21-25	8.5	12.0	28.7	40.0	17.0	2.4		
	26-40	26.2	12.1	28.2	36.5	20.3	2.9		
	41-55	30.5	8.9	26.4	36.7	25.9	2.2		
	56-70	19.4	3.6	16.7	36.1	39.6	4.0		
	71+	8.0	0.8	7.2	25.0	52.2	14.8		
Sex	Female	47.5	7.7	21.1	34.9	32.1	4.2	3	187
	Male	52.5	8.4	24.6	36.2	26.9	4.0		
Geographic area	Rural	16.6	8.6	24.2	35.5	28.2	3.6	3	20
	Urban	83.4	7.9	22.7	35.6	29.6	4.2		
Immigrant	No	74.5	8.7	23.2	37.6	26.3	4.2	3	702
	Yes	25.5	6.2	22.1	29.6	38.5	3.6		
Race	White	77.8	8.4	23.5	37.9	27.9	2.3	6	5,570
	Non-White	16.5	6.2	23.1	28.9	38.5	3.4		
	Missing	5.6	8.6	15.3	22.5	23.1	30.6		
Marital status	Married or common-law	65.8	7.9	23.1	36.0	29.5	3.6	3	64
	Not married	34.3	8.3	22.7	34.8	29.2	5.0		
Education	Less than secondary	4.4	4.7	13.5	28.9	49.0	4.0	12	5,166
	Secondary graduate	9.5	5.8	21.8	32.9	36.5	3.1		
	Some post-secondary	4.7	8.7	21.2	35.5	32.6	2.1		
	Post-secondary graduate	73.5	8.6	24.4	37.6	27.4	2.1		
	Missing	7.9	7.1	17.3	23.9	27.0	24.8		
Income	$0-$19,999	2.9	5.8	18.0	27.8	44.9	3.4	15	2,197
	$20,000-$39,999	7.9	5.8	18.6	30.0	41.8	3.8		
	$40,000-$59,999	9.9	7.0	22.0	35.1	33.5	2.4		
	$60,000-$79,999	10.5	8.2	25.8	35.8	28.3	1.8		
	$80,000 or more	33.6	10.7	26.9	38.9	22.1	1.4		
	Missing	35.3	6.4	19.9	34.3	31.5	7.9		
Employment	Employed	76.3	10.0	26.7	37.5	24.3	1.6	3	5,452
	Unemployed	23.7	1.9	11.0	29.2	45.9	12.1		

Descriptive characteristic		Speeding							
		N (%)	Much faster	Little faster	The same	Little slower	Much slower	*df*	χ2
Self-perceived health	Healthy	89.5	8.0	23.5	36.6	28.8	3.1	3	1,134
	Unhealthy	10.5	7.9	18.3	26.6	34.9	12.2		
Mental health	Positive	93.1	7.8	23.5	36.8	30.1	1.8	3	8,490
	Negative	6.9	10.6	15.6	19.1	20.1	34.6		
Stress	Lower levels of stress	32.6	4.3	17.0	36.3	37.9	4.4	3	1,307
	Higher levels of stress	67.4	9.8	25.8	35.2	25.3	3.9		
Mood	No mood	93.5	7.8	22.9	35.8	29.6	3.9	3	129
	Mood	6.5	11.3	23.6	31.9	26.4	6.8		
Anxiety	No anxiety	95.4	7.9	22.9	35.8	29.5	3.9	3	89
	Anxiety	4.6	11.3	23.3	31.3	27.2	6.9		
Smoking	Non-smoker	79.6	7.3	22.8	35.8	29.7	4.5	3	229
	Smoker	20.4	11.1	23.7	34.7	28.1	2.5		
Alcohol	No alcohol	20.0	5.4	16.2	26.0	38.1	14.3	3	3,965
	Alcohol	80.0	8.7	24.6	38.0	27.2	1.5		
Binge drinking	Non-drinker	20.4	5.4	16.2	26.2	37.9	14.3	6	5,158
	Drinker, non-binger	40.1	6.5	21.0	36.3	34.0	2.2		
	Binger	39.5	10.9	28.4	39.6	20.4	0.7		
RWDD	No	89.9	7.4	21.9	35.4	30.8	4.5	3	856
	Yes	10.1	13.4	31.9	37.5	17.0	0.2		
Injuries	Not injured	85.4	7.5	22.6	35.1	30.5	4.3	6	312
	1 injury	10.9	10.9	23.6	39.7	23.3	2.4		
	2 or more	3.7	11.9	27.9	35.2	21.8	3.2		
Continuous variables									*t*
Age2	*M*[a]	44.8	38.6	41.4	44.0	49.2	52.5		571
	SE[b]	0.1	12.5	14.1	16.2	18.9	23.6		
Satisfaction with life	*M*	8.1	7.8	7.9	8.1	8.1	9.1		1,087
	SE	0.0	1.8	1.5	1.5	1.7	1.6		

[a] Mean, [b] Standard error of the mean

Table 6
Percent distribution of driver characteristics across aggressive driving ($p < .05$)

Characteristic		Aggressive Driving								
		N (%)	Much more	Little more	The same	Little less	Much less	Total	*df*	χ2
Age	<=20	7.4	5.2	16.8	32.8	38.1	7.2	100	15	4,970
	21-25	8.5	12.0	28.6	40.0	16.9	2.5	100		
	26-40	26.2	12.1	28.2	36.4	20.3	3.0	100		
	41-55	30.5	8.8	26.3	36.6	25.6	2.7	100		
	56-70	19.4	3.6	16.7	35.9	39.3	4.6	100		
	71+	8.0	0.8	7.0	24.8	51.4	16.0	100		
Sex	Female	47.5	7.7	21.1	34.8	31.8	4.7	100	3	182
	Male	52.5	8.3	24.6	36.1	26.7	4.4	100		
Geographic area	Rural	16.6	8.5	24.2	35.4	27.7	4.3	100	3	18
	Urban	83.4	7.9	22.6	35.5	29.4	4.6	100		
Immigrant	No	74.5	8.6	23.2	37.5	26.0	4.7	100	3	704
	Yes	25.5	6.2	22.0	29.5	38.2	4.2	100		
Race	White	77.8	8.4	23.4	37.8	27.7	2.7	100	6	5,370
	Non-White	16.5	6.2	22.9	28.9	38.2	3.8	100		
	Missing	5.6	8.6	15.3	22.3	22.1	31.8	100		
Marital Status	Married or common-law	65.8	7.9	23.0	35.9	29.3	4.0	100	3	68
	Not married	34.3	8.3	22.7	34.7	28.8	5.6	100		
Education	Less than secondary	4.4	4.7	13.4	28.7	48.2	4.9	100	12	5,056
	Secondary graduate	9.5	5.8	21.8	32.8	36.2	3.5	100		
	Some post-secondary	4.7	8.7	21.1	35.3	32.5	2.4	100		
	Post-secondary graduate	73.5	8.6	24.3	37.5	27.2	2.5	100		
	Missing	7.9	7.0	17.1	23.7	26.1	26.1	100		
Income	$0-$19,999	2.9	5.8	18.0	27.7	44.6	3.9	100	15	2,304
	$20,000-$39,999	7.9	5.8	18.6	29.9	41.4	4.3	100		
	$40,000-$59,999	9.9	7.0	22.0	35.0	33.4	2.6	100		
	$60,000-$79,999	10.5	8.2	25.8	35.8	28.3	1.9	100		
	$80,000 or more	33.6	10.7	26.9	38.9	22.0	1.6	100		
	Missing	35.3	6.4	19.8	34.2	30.9	8.7	100		
Employment status	Employed	76.3	9.9	26.6	37.5	24.1	1.9	100	3	5,366
	Unemployed	23.7	1.9	10.9	29.1	45.3	12.9	100		
Self-perceived health	Healthy	89.5	8.0	23.4	36.5	28.5	3.5	100	3	1,097
	Unhealthy	10.5	7.9	18.2	26.5	34.5	12.9	100		
Mental health	Positive	93.1	7.8	23.5	36.7	29.8	2.3	100	3	7,833
	Negative	6.9	10.6	15.3	19.0	19.9	35.4	100		

Characteristic		Aggressive Driving								
		N (%)	Much more	Little more	The same	Little less	Much less	Total	*df*	χ2
Stress	Lower levels of stress	32.6	4.3	17.0	36.2	37.5	5.0	100	3	1,305
	Higher levels of stress	67.4	9.8	25.7	35.1	25.1	4.3	100		
Mood	No mood	93.5	7.8	22.8	35.7	29.3	4.4	100	3	109
	Mood	6.5	11.3	23.6	31.9	26.3	6.9	100		
Anxiety	No anxiety	95.4	7.9	22.9	35.7	29.2	4.4	100	3	81
	Anxiety	4.6	11.2	23.3	31.3	26.9	7.2	100		
Smoker	Non-smoker	79.6	7.2	22.7	35.7	29.4	5.0	100	3	226
	Smoker	20.4	11.0	23.7	34.6	27.9	2.8	100		
Alcohol	No alcohol	20.0	5.4	16.2	25.9	37.4	15.1	100	3	3,873
	Alcohol	80.0	8.7	24.6	37.9	27.0	1.9	100		
Binge drinking	Non-drinker	20.4	5.4	16.1	26.1	37.3	15.2	100	6	5,102
	Drinker non-binger	40.1	6.5	20.9	36.2	33.7	2.7	100		
	Binger	39.5	10.9	28.4	39.5	20.3	0.9	100		
RWDD	No RWDD	89.9	7.4	21.9	35.3	30.5	5.0	100	3	861
	RWDD	10.1	13.4	31.9	37.4	16.9	0.4	100		
Number of injuries	Not injured	85.4	7.5	22.6	35.0	30.2	4.8	100	6	308
	1 time	10.9	10.9	23.6	39.5	23.2	2.8	100		
	2 or more times	3.7	11.9	27.9	35.1	21.5	3.6	100		
Continuous measures										*t*
Age2	*M*[a]	44.8	38.6	41.4	44.0	49.2	52.6	47,355		571
	SE[b]	0.1	12.5	14.1	16.1	18.8	23.2			
Satisfaction with life	*M*	8.1	7.8	7.9	8.1	8.1	9.0	47,355		1,087
	SE	0.0	1.8	1.5	1.5	1.7	1.7			

[a] Mean, [b] Standard error of the mean.

Table 7
Percent distribution of driver characteristics across driving under the influence of alcohol (DUIA) ($p < .05$)

Characteristic		DUIA					
		N (%)	Yes	No	Total	*df*	χ2
Age	<=20	7.4	2.6	97.4	100	15	317
	21-25	8.5	8.8	91.2	100		
	26-40	26.2	6.6	93.4	100		
	41-55	30.5	5.6	94.4	100		
	56-70	19.4	3.8	96.2	100		
	71+	8.0	2.0	98.0	100		
Sex	Female	47.5	1.6	98.4	100	3	1,167
	Male	52.5	8.6	91.4	100		
Geographic area	Rural	16.6	6.2	93.8	100	3	17
	Urban	83.4	5.1	94.9	100		
Immigrant	No	74.5	5.9	94.1	100	3	108
	Yes	25.5	3.4	96.6	100		
Race	White	77.8	5.9	94.1	100	6	130
	Non-White	16.5	3.0	97.0	100		
	Missing	5.6	3.3	96.7	100		
Marital Status	Married or common-law	65.8	4.9	95.1	100	3	27
	Not married	34.3	6.0	94.0	100		
Education	Less than secondary	4.4	3.9	96.1	100	12	29
	Secondary graduate	9.5	5.5	94.5	100		
	Some post-secondary	4.7	6.6	93.4	100		
	Post-secondary graduate	73.5	5.4	94.6	100		
	Missing	7.9	4.0	96.0	100		
Income	$0-$19,999	2.9	3.8	96.2	100	15	318
	$20,000-$39,999	7.9	3.8	96.2	100		
	$40,000-$59,999	9.9	4.8	95.2	100		
	$60,000-$79,999	10.5	5.9	94.1	100		
	$80,000 or more	33.6	7.6	92.4	100		
	Missing	35.3	3.4	96.6	100		
Employment	Employed	76.3	6.2	93.8	100	3	269

Characteristic		DUIA					
		N (%)	Yes	No	Total	*df*	χ2
	Unemployed	23.7	2.3	97.8	100		
Self-perceived health	Healthy	89.5	5.4	94.6	100	3	22
	Unhealthy	10.5	3.9	96.2	100		
Mental health	Positive	93.1	5.4	94.6	100	3	29
	Negative	6.9	3.2	96.8	100		
Stress	Lower levels of stress	32.6	4.8	95.2	100	3	9
	Higher levels of stress	67.4	5.5	94.5	100		
Mood	No mood	93.5	5.3	94.7	100	3	4
	Mood	6.5	4.5	95.5	100		
Anxiety	No anxiety	95.4	5.3	94.7	100	3	0.4
	Anxiety	4.6	5.0	95.0	100		
Smoking	Non-smoker	79.6	4.5	95.5	100	3	221
	Smoker	20.4	8.3	91.7	100		
Alcohol	No alcohol	20.0	0.0	100.0	100	3	657
	Alcohol	80.0	6.6	93.4	100		
Binge drinking	Non-drinker	20.4	0.1	99.9	100	6	2,740
	Drinker, non-binger	40.1	1.4	98.7	100		
	Binger	39.5	11.9	88.1	100		
RWDD	No	89.9	2.5	97.5	100	3	6,383
	Yes	10.1	29.8	70.2	100		
Injuries	Not injured	85.4	4.9	95.1	100	6	93
	1 injury	10.9	7.1	92.9	100		
	2 or more injuries	3.7	8.9	91.1	100		
Continuous measures							*t*
Age2	*M*[a]	44.8	41.2	45.0		47,355	571
	SE[b]	0.1	14.6	17.2			
Satisfaction with life	*M*	8.1	7.9	8.1		47,355	1,087
	SE	0.0	1.6	1.6			

[a] Mean, [b] standard error of the mean.

Table 8
K-means five-cluster solution

RDB	*N* (%)	Average Drivers[a] 14,179 (30.0)	Cautious Drivers[b] 15,791 (33.3)	Beltless Drivers[c] 2,083 (4.4)	Egocentric Drivers[d] 5,558 (11.7)	Poly-risk Drivers[e] 9,745 (20.6)	*F* Score[f]
Seat belt◆							
M[g] (*SD*[h])	3.81 (0.66)	3.95 (0.24)	3.97 (0.16)	1.04 (0.19)	3.84 (0.48)	3.92 (0.36)	46,288*
CI[i] (99%)	(3.80, 3.82)	(3.95, 3.95)	(3.97, 3.98)	(1.03, 1.05)	(3.83, 3.85)	(3.91, 3.93)	
Cell Phone							
M (*SD*)	3.34 (0.95)	3.66 (0.47)	3.66 (0.72)	3.94 (0.36)	1.59 (0.49)	3.58 (0.59)	19,807*
CI (99%)	(3.33, 3.35)	(3.65, 3.67)	(3.64, 3.67)	(3.92, 4.00)	(1.58, 1.60)	(3.58, 3.60)	
Fatigue◆							
M (*SD*)	2.94 (0.94)	3.00 (0.04)	4.00 (0.06)	3.98 (0.14)	2.32 (0.67)	1.72 (0.45)	78,490*
CI (99%)	(2.93, 2.95)	(3.00, 3.00)	(4.00, 4.00)	(3.97, 3.99)	(2.30, 2.33)	(1.71, 1.73)	
Speeding◆							
M (*SD*)	2.99 (1.00)	3.00(0.04)	4.00 (0.09)	4.88 (0.38)	2.32 (0.6)	1.72 (0.45)	85,333*
CI (99%)	(2.98, 3.00)	(3.00, 3.00)	(4.00, 4.00)	(4.87, 4.90)	(2.30, 2.33)	(1.71, 1.73)	
Aggressive Driving◆							
M (*SD*)	2.99 (1.01)	3.00 (0.08)	4.02 (0.13)	4.88 (0.37)	2.32 (0.67)	1.72 (0.46)	81,991*
CI (99%)	(2.98, 3.00)	(3.00, 3.00)	(4.02, 4.02)	(4.87, 4.90)	(2.30, 2.34)	(1.71, 1.73)	
DUIA							
M (*SD*)	0.05 (0.22)	0.04 (0.02)	0.02 (0.14)	0.01 (0.01)	0.13 (0.34)	0.07 (0.25)	328*
CI (99%)	(0.05, 0.05)	(0.04, 0.04)	(0.02, 0.02)	(0.01, 0.01)	(0.12, 0.14)	(0.06, 0.07)	

[a] Cluster 1, [b] Cluster 2, [c] Cluster 3, [d] Cluster 4, [e] Cluster 5, [f] *F* score generated by ANOVA tests evaluated between group differences in risky driving behaviours in the *K*-means cluster solution, *df* (4, 47351), [g] mean, [h] standard deviation, [i] confidence interval, *$p<.001$□
□Post hoc Bonferroni tests [*df* (4, 47351), *p*<.001] confirmed pairwise comparisons were significantly different between all five clusters.

Table 9
Percent distributions of sociodemographic variables by cluster

Sociodemographic measures		*N* %	Average Drivers	Cautious Drivers	Beltless Drivers	Egocentric Drivers	Poly-risk Drivers
Age	<=20	7.4	6.6	9.7	12.3	6.9	4.9
	21-25	8.5	8.2	4.9	4.9	16.1	8.9
	26-40	26.2	24.9	18.2	18.1	37.3	32.6
	41-55	30.5	31.9	26.8	17.7	30.5	36.0
	56-70	19.4	21.7	26.2	18.8	8.9	14.8
	71+	8.0	6.8	14.3	28.2	0.3	2.8
Sex	Female	47.5	48.7	52.2	47.1	37.2	47.2
	Male	52.5	51.3	47.8	52.9	62.8	52.9
Geographic Area	Rural	16.6	17.2	15.8	15.6	14.8	18.3
	Urban	83.4	82.8	84.2	84.4	85.2	81.7
Immigrant Status	Non-immigrant	77.9	66.6	77.6	80.2	76.2	74.5
	Immigrant	22.1	33.4	22.4	19.8	23.8	25.5
Race	White	77.8	82.5	73.9	46.9	81.5	80.5
	Non-White	16.5	14.0	21.6	13.3	14.8	14.9
	Missing	5.6	3.5	4.5	39.9	3.8	4.6
Marital Status	Married or common-law	65.8	67.3	66.2	57.2	61.9	67.6
	Not married	34.3	32.7	33.9	42.8	38.2	32.4
Education	Less than secondary	4.4	3.9	7.2	5.4	2.1	2.7
	Secondary school graduate	9.5	9.0	11.8	7.8	8.9	8.1
	Some post-secondary education	4.7	4.7	5.2	2.6	4.8	4.5
	Post-secondary school graduate	73.5	77.3	68.5	39.2	78.2	78.6
	Missing	7.9	5.2	7.3	45.0	6.0	6.2
Income	$0-$19,999	2.9	2.4	4.4	3.0	1.6	2.4
	$20,000-$39,999	7.9	7.3	11.1	8.5	4.5	6.8
	$40,000-$59,999	9.9	10.2	11.3	5.9	8.2	9.7
	$60,000-$79,999	10.5	10.5	10.1	4.3	10.7	11.9
	$80,000 or more	33.6	35.4	25.3	12.2	43.7	39.2
	Missing	35.3	34.3	37.9	66.1	31.3	30.0
Employment Status	Employed	76.3	77.8	63.1	31.6	93.4	88.7
	Unemployed	23.7	22.2	36.9	68.4	6.6	11.3
Age2	M[a] (*SE*[b])	44.8(0.1)	49.5(0.2)	54.8(0.2)	55.1(0.5)	37.3(0.2)	44.2(0.2)

[a] Mean, [b] Standard error of the mean

Table 10
Percent distributions of health-related variables in the sample population and by cluster

Health-related measures		*N* %	Average Drivers	Cautious Drivers	Beltless Drivers	Egocentric Drivers	Poly-risk Drivers
Self-Perceived Health	Healthy	89.5	91.7	87.6	68.7	94.2	89.9
	Unhealthy	10.5	8.4	12.4	31.3	5.8	10.1
Mental Health Status	Positive	93.1	96.0	95.2	45.1	96.6	93.6
	Negative	6.9	4.0	4.8	54.9	3.4	6.4
Stress	Lower levels of stress	32.6	34.0	41.9	35.8	24.6	23.4
	Higher levels of stress	67.4	66.0	58.1	64.2	75.4	76.6
Mood Disorder	No mood	93.5	93.8	94.2	89.5	94.8	92.0
	Mood disorder	6.5	6.2	5.9	10.5	5.2	8.0
Anxiety Disorder	No anxiety	95.4	95.7	95.8	92.5	95.8	94.8
	Anxiety	4.6	4.3	4.3	7.5	4.2	5.2
Smoking	Non-smoker	79.6	80.5	80.7	85.0	77.1	77.5
	Smoker	20.4	19.5	19.3	15.0	22.9	22.5
Alcohol	Does not consume alcohol	20.0	15.5	25.9	67.5	10.6	15.2
	Consumes alcohol	80.0	84.5	74.2	32.5	89.4	84.8
Continuous measures M^a (SE^b)							
Satisfaction With Life		8.1 (1.6)	8.1 (0.01)	8.1 (0.01)	9.0 (0.4)	8.0 (0.2)	7.8 (0.2)

[a] Mean, [b] Standard error of the mean

Table 11
Percent distributions of other risky behaviours by cluster

Other risky behaviours		*N* %	Average Drivers	Cautious Drivers	Beltless Drivers	Egocentric Drivers	Poly-risk Drivers
Binge Drinking	Does not drink	20.4	15.9	26.3	68.6	11.0	15.5
	Drinks, but does not binge	40.1	44.0	46.4	22.4	28.9	38.0
	Binger	39.5	40.2	27.3	9.0	60.2	46.5
RWDD	Does not RWDD	89.9	90.9	94.3	98.3	79.8	88.3
	RWDD	10.1	9.1	5.7	1.7	20.2	11.7
Number of Injuries	Not injured	85.4	84.3	88.6	90.1	83.0	83.3
	1 injury	10.9	12.2	8.6	7.0	11.8	12.2
	2 or more	3.7	3.4	2.8	3.0	5.2	4.5

Table 12
ANOVA results comparing sociodemographic variables in the *k*-mean five-cluster solution

Variable (Code Range)	*N*	Average Drivers[a]	Cautious Drivers[b]	Beltless Drivers[c]	Egocentric Drivers[d]	Poly-Risk Drivers[e]	*F* Score[f]
n		14,179	15,791	2,083	5,558	9,745	
Age[•] (1-6)							
M (*SD*)[g]	3.70 (1.30)	3.74 (1.24)	3.98 (1.35)	4.10 (1.66)	3.19 (1.20)	3.55 (1.13)	544*
CI (99%)[h]	(3.68, 3.72)	(3.72, 3.76)	(3.95, 4.00)	(4.03, 4.17)	(3.17, 3.57)	(3.53, 3.57)	
Sex (0-1)	47,356	14,179	15,791	2,083	5,558	9,745	
M (*SD*)	0.52 (0.50)	0.51 (0.49)	0.48 (0.47)	0.53 (0.50)	0.63 (0.55)	0.53 (0.52)	112*
CI (99%)	(0.51, 0.53)	(0.50, 0.52)	(0.47, 0.49)	(0.51, 0.55)	(0.61, 0.64)	(0.52, 0.54)	
Geographic area (0-1)	47,356	14,179	15,791	2,083	5,558	9,745	
M (*SD*)	0.83 (0.37)	0.83 (0.37)	0.84 (0.34)	0.84 (0.36)	0.85 (0.41)	0.82 (0.40)	13*
CI (99%)	(0.83, 0.83)	(0.82, 0.83)	(0.84, 0.85)	(0.83, 0.86)	(0.84, 0.86)	(0.81, 0.82)	
Immigrant status (0-1)	47,356	14,179	15,791	2,083	5,558	9,745	
M (*SD*)	0.25 (0.44)	0.22 (0.41)	0.33 (0.44)	0.22 (0.41)	0.20 (0.46)	0.24 (0.44)	175*
CI (99%)	(0.24, 0.26)	(0.21, 0.23)	(0.33, 0.34)	(0.21, 0.24)	(0.19, 0.21)	(0.23, 0.25)	
Race (0-1)	44,334	13,060	13,154	1,222	6,976	9,923	
M (*SD*)	1.18 (0.38)	1.15 (0.35)	1.23 (0.40)	1.22 (0.52)	1.15 (0.42)	1.16 (0.39)	98*
CI (99%)	(1.18, 1.18)	(1.14, 1.15)	(1.22, 1.23)	(1.20, 1.24)	(1.14, 1.16)	(1.15, 1.16)	
Marital status (0-1)	47,356	14,179	15,791	2,083	5,558	9,745	
M (*SD*)	0.34 (0.47)	0.33 (0.46)	0.34 (0.44)	0.43 (0.49)	0.38 (0.56)	0.32 (0.49)	37*
CI (99%)	(0.32, 0.34)	(0.32, 0.33)	(0.33, 0.35)	(0.41, 0.45)	(0.37, 0.39)	(0.31, 0.33)	
Education (0-4)	44,201	13,097	13,036	1,142	6,961	9,965	
M (*SD*)	3.60 (0.86)	3.64 (0.82)	3.46 (0.95)	3.38 (1.34)	3.69 (0.86)	3.69 (0.79)	170*
CI (99%)	(3.59, 3.61)	(3.62, 3.65)	(3.44, 3.47)	(3.32, 3.43)	(3.67, 3.71)	(3.68, 3.71)	

Variable (Code Range)	*N*	Average Drivers[a]	Cautious Drivers[b]	Beltless Drivers[c]	Egocentric Drivers[d]	Poly-Risk Drivers[e]	*F* Score[f]
n		14,179	15,791	2,083	5,558	9,745	
Income◆ (0-5)	31,005	9,070	8,726	704	5,076	7,429	
M (*SD*)	3.99 (1.25)	4.05 (1.44)	3.66 (1.60)	3.42 (2.26)	4.32 (1.46)	4.12 (1.46)	317*
CI (99%)	(3.97, 4.01)	(4.02, 4.08)	(3.62,3.69)	(3.29, 3.55)	(4.27, 4.36)	(4.09, 4.16)	
Employment status◆ (0-1)	47,356	14,179	15,791	2,083	5,558	9,745	
M (*SD*)	0.24 (0.43)	0.22 (0.41)	0.37 (0.45)	0.68 (0.46)	0.07 (0.28)	0.11 (0.33)	1,621*
CI (99%)	(0.23, 0.25)	(0.22, 0.23)	(0.36, 0.38)	(0.66, 0.70)	(0.06, 0.07)	(0.11, 0.12)	
Age2◆	47,356	14,179	15,791	2,083	5,558	9,745	
M (*SE*)[i]	44.84 (0.01)	49.49*(0.15)	54.81*(0.15)	55.07*(0.53)	37.33*(0.18)	44.17*(0.16)	805*
CI (99%)	(44.64, 45.04)	(49.11, 49.87)	(54.40, 55.21)	(53.78, 56.50)	(36.87, 37.79)	(43.76, 44.58)	

[a] Cluster 1, [b] Cluster 2, [c] Cluster 3, [d] Cluster 4, [e] Cluster 5, [f] *df* (4, 47351), $p<.001$] [g] mean (standard deviation), [h] confidence interval.
*$p<0.01$, Post hoc Bonferroni tests demonstrated significant differences between all five clusters for age [*df* (4, 47351), $p<.001$], income [*df* (4, 3100), $p<.001$)], and employment [*df* (4, 47351), $p<.001$].

Table 13
ANOVA results comparing health-related variables in the *k*-mean five-cluster solution

Variable (Code range)	*N*	Average Drivers[a]	Cautious Drivers[b]	Beltless Drivers[c]	Egocentric Drivers[d]	Poly-Risk Drivers[e]	*F* Score[f]
n	47,356	14,179	15,791	2,083	5,558	9,745	
Self-perceived Health• (0-1)							
M (*SD*)[g]	0.11 (0.31)	0.08 (0.27)	0.12 (0.31)	0.31 (0.46)	0.06 (0.27)	0.10 (0.31)	317.84*
CI (99%)[h]	(0.11, 0.11)	(0.08, 0.09)	(0.12, 0.13)	(0.29, 0.33)	(0.05, 0.06)	(0.09, 0.11)	
Mental Health Status (0-1)							
M (*SD*)	0.07 (0.25)	0.04 (0.19)	0.05 (0.20)	0.55 (0.49)	0.03 (0.21)	0.06 (0.25)	2,331.96*
CI (99%)	(0.07, 0.07)	(0.04, 0.04)	(0.04, 0.05)	(0.52, 0.58)	(0.03, 0.04)	(0.06, 0.07)	
Stress (0-1)							
M (*SD*)	0.67 (0.47)	0.66 (0.46)	0.58 (0.46)	0.64 (0.48)	0.75 (0.49)	0.77 (0.44)	305.40*
CI (99%)	(0.66, 0.68)	(0.65, 0.67)	(0.57, 0.59)	(0.62, 0.67)	(0.74, 0.77)	(0.76, 0.78)	
Mood Disorder (0-1)							
M (*SD*)	0.07 (0.25)	0.06 (0.24)	0.06 (0.22)	0.11 (0.30)	0.05 (0.25)	0.08 (0.28)	30.90*
CI (99%)	(0.07, 0.07)	(0.06, 0.07)	(0.05, 0.06)	(0.09, 0.12)	(0.04, 0.06)	(0.07, 0.09)	
Anxiety Disorder (0-1)							
M (*SD*)	0.05 (0.21)	0.04 (0.20)	0.04 (0.19)	0.07 (0.26)	0.04 (0.23)	0.05 (0.23)	14.16*
CI (99%)	(0.05, 0.05)	(0.04, 0.05)	(0.04, 0.05)	(0.06, 0.09)	(0.04, 0.05)	(0.05, 0.06)	
Smoking (0-1)							
M (*SD*)	0.20 (0.40)	0.20 (0.39)	0.19 (0.37)	0.15 (0.35)	0.23 (0.48)	0.22 (0.43)	27.38*
CI (99%)	(0.20, 0.20)	(0.19, 0.20)	(0.19, 0.20)	(0.13, 0.17)	(0.21, 0.24)	(0.21, 0.24)	
Alcohol (0-1)							
M (*SD*)	0.80 (0.40)	0.84 (0.36)	0.74 (0.41)	0.33 (0.46)	0.89 (0.35)	0.85 (0.37)	1,006.38*
CI (99%)	(0.80, 0.80)	(0.84, 0.85)	(0.73, 0.75)	(0.30, 0.35)	(0.88, 0.90)	(0.84, 0.86)	
Satisfaction With Life (1-10)							
M (*SE*)[i]	8.06 (0.01)	8.11 (0.01)	8.07 (0.1)	9.01 (0.04)	8.04 (0.02)	7.82 (0.02)	242.04*
CI (99%)	(8.04, 8.08)	(8.07, 8.13)	(8.04, 8.11)	(8.92, 9.11)	(7.99, 8.09)	(7.78, 7.87)	

[a] Cluster 1, [b] Cluster 2, [c] Cluster 3, [d] Cluster 4, [e] Cluster 5, [f] *df* (4, 47351), p<.001] [g] mean (standard deviation), [h] confidence interval, [i] standard error of the mean.
*p<0.01, •Post hoc Bonferroni tests showed significant differences between all five clusters for self-perceived health [*df* (4, 47351), p<.001].

Table 14
ANOVA results comparing other risk-taking variables in the *k*-mean five-cluster solution

Variable (Code Range)	*N*	Average Drivers[a]	Cautious Drivers[b]	Beltless Drivers[c]	Egocentric Drivers[d]	Poly-Risk Drivers[e]	*F* Score[f]
n	47,356	14,179	15,791	2,083	5,558	9,745	
Binge Drinking• (1-3)							
M (*SD*)[g]	2.19 (0.75)	2.24 (0.69)	2.01 (0.69)	1.40 (0.64)	2.49 (0.79)	2.31 (0.75)	1,266.02*
CI (99%)[h]	(2.18, 2.20)	(2.23, 2.25)	(2.00, 2.02)	(1.39, 1.41)	(2.48, 2.50)	(2.30, 2.32)	
RWDD• (0-1)							
M (*SD*)	0.10 (0.30)	0.09 (0.28)	0.06 (0.22)	0.02 (0.13)	0.20 (0.46)	0.12 (0.33)	341.73*
CI (99%)	(0.10, 0.10)	(0.08, 0.10)	(0.05, 0.06)	(0.01, 0.02)	(0.19, 0.22)	(0.11, 0.12)	
Number of Injuries (1-3)							
M (*SD*)	1.18 (0.47)	1.19 (0.46)	1.14 (0.39)	1.13 (0.41)	1.22 (0.60)	1.21 (0.53)	57.72*
CI (99%)	(2.81, 2.83)	(2.80, 2.82)	(2.85, 2.87)	(2.85, 2.89)	(2.76, 2.80)	(2.77, 2.80)	

[a] Cluster 1, [b] Cluster 2, [c] Cluster 3, [d] Cluster 4, [e] Cluster 5, [f] *df* (4, 47351), $p<.001$, [g] mean (standard deviation), [h] confidence interval.
*$p<0.01$, •Bonferroni post-hoc tests demonstrated significant differences in the mean scores for binge drinking and RWDD for all five clusters [*df* (4, 47351), $p<.001$].

4.7 References

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster Analysis*. Newberry Park, California: Sage University paper series on Quantitative Applications Social Science in the Social Sciences. 07-044.

Anderson, L. R., & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. Journal of Health Economics, 27(5), 1260-1274.

Asbridge, M., Brubacher, J. R., & Chan, H. (2013). Cell phone use and traffic crash risk: A culpability analysis. International Journal of Epidemiology, 42(1), 259-267.

Baird, B. M., Lucas, R. E., & Donnellan, M. B. (2010). Life satisfaction across the lifespan: Findings from two nationally representative panel studies. Social Indicators Research, 99(2), 183-203.

Bhatti, J. A., Constant, A., Salmi, L. R., Chiron, M., Lafont, S., Zins, M., & Lagarde, E. (2008). Impact of retirement on risky driving behavior and attitudes towards road safety among a large cohort of French drivers (the GAZEL cohort). Scandinavian Journal of Work, Environment & Health, 34(4), 307.

Bible, J., Datta, S., & Datta, S. (2013). Cluster Analysis: Finding groups in data. Informatics for Materials Science and Engineering: Data-Driven Discovery for Accelerated Experimentation and Application, 53.

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(4), 1282-1293.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics-Theory and Methods, 3(1), 1-27.

CCMTA – STRID sub-group on fatigue. Retrieved from http://www.ccmta.ca/english/committees/rsrp/strid-fatigue/strid-fatigue-reports.cfm

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., & Newstead, S. (2003). Self-regulatory behaviours of older drivers. Annual proceedings/Association for the Advancement of Automotive Medicine, Association for the Advancement of Automotive Medicine. 47 181.

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., Newstead, S., Koppel, S., & O'Hare, M. (2006). Characteristics of older drivers who adopt self-regulatory driving behaviours. Transportation Research Part F: Traffic Psychology and Behaviour, 9(5), 363.

Dahlen, E. R., Martin, R. C., Ragan, K., & Kuhlman, M. M. (2005). Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving. Accident Analysis & Prevention, 37(2), 341-348.

Deery, H. A., & Fildes, B. N. (1999). Young novice driver subtypes: Relationship to high-risk behavior, traffic accident record, and simulator driving performance. Human Factors: The Journal of the Human Factors and Ergonomics Society, 41(4), 628.

Di Milia, L., Smolensky, M. H., Costa, G., Howarth, H. D., Ohayon, M. M., & Philip, P. (2011). Demographic factors, fatigue, and driving accidents: An examination of the published literature. Accident Analysis & Prevention, 43(2), 516-532.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the European Economic Association, 9(3), 522-550.

Duda, R. O., & Hart, P. E. (1973). Pattern classification and scene analysis Wiley New York.

Duda, R. O., Hart, P. E., & Stork, D. G. (2012). Pattern classification John Wiley & Sons.

Evans, L. (1991). Traffic safety and the driver. New York: Van Nostrand Reinhold.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). In Shewhart W. A., Wilks S. S. (Eds.), Cluster analysis, 5th edition, John Wiley & Sons, Ltd, Chichester, UK. (5th ed.). doi:10.1002/9780470977811

Fernandes, R., Hatfield, J., & Job, R. (2010). A systematic investigation of the differential predictors for speeding, drink-driving, driving while fatigued, and not wearing a seat belt, among young drivers. Transportation Research Part F: Traffic Psychology and Behaviour, 13(3), 179-196.

Gielen, A. C., & Sleet, D. (2003). Application of behavior-change theories and methods to injury prevention. Epidemiologic Reviews, 25, 65-76.

Golias, I., & Karlaftis, M. G. (2001). An international comparative study of self-reported driver behavior. Transportation Research Part F: Traffic Psychology and Behaviour, 4(4), 243-256.

Green, L. W. & Kreuter, M. W. (1999). *Health promotion planning: An educational and environmental approach* (3rd ed.). Mountain View, CA: Mayfield Press.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. Journal of Health & Social Behavior, 38(1), 21-37.

Jessor, R. (1987). Risky driving and adolescent problem behaviour: Theoretical and empirical linkage. T Benjamin (Ed.), Young Drivers Impaired by Alcohol and Other Drugs, Royal Society of Medicine Services, London, 97-110.

Jessor, R., Turbin, M. S., & Costa, F. M. (1997). Predicting developmental change in risky driving: The transition to young adulthood. Applied Developmental Science, 1(1), 4.

Jessor, R. (1987). Risky driving and adolescent problem behavior: An extension of problem-behavior theory. Alcohol, Drugs & Driving, 3(3-4), 1-11.

Jonah, B. A., & Dawson, N. E. (1987). Youth and risk: Age differences in risky driving, risk perception, and risk utility. Alcohol, Drugs & Driving.

Juster, F. T., & Smith, J. P. (1997). Improving the quality of economic data: Lessons from the HRS and AHEAD. Journal of the American Statistical Association, 92(440), 1268-1278.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. Archives of General Psychiatry, 62(6), 593-602.

Kessler, R. C., Olfson, M., & Berglund, P. A. (1998). Patterns and predictors of treatment contact after first onset of psychiatric disorders. American Journal of Psychiatry, 155(1), 62-69.

Lajunen, T., & Summala, H. (2003). Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. Transportation Research Part F: Traffic Psychology and Behaviour, 6(2), 97-107.

Lajunen, T., & Parker, D. (2001). Are aggressive people aggressive drivers? A study of the relationship between self-reported general aggressiveness, driver anger and aggressive driving. Accident Analysis & Prevention, 33(2), 243-255.

Linden, R., Mann, R. E., Smart, R. G., Vingilis, E., Solomon, R., Chamberlain, E., . . . Stoduto, G. (2010). Research, policy development, and progress: Antisocial behaviour and the automobile. Canadian Public Policy, 36, 81-93.

Lucidi, F., Giannini, A. M., Sgalla, R., Mallia, L., Devoto, A., & Reichmann, S. (2010). Young novice driver subtypes: Relationship to driving violations, errors and lapses. Accident Analysis & Prevention, 42(6), 1689-1696.

Lustman, M., Wiesenthal, D. L., & Flett, G. L. (2010). Narcissism and aggressive driving: Is an inflated view of the self a road hazard? Journal of Applied Social Psychology, 40, 1423-1449.

Mann, R. E., Asbridge, M., Stoduto, G., Smart, R. G., Goldbloom, D. S., Vingilis, E. R., & Wickens, C. M. (2010). Psychological distress and collision involvement among adult drivers. Stress and Health, 26(2), 127.

Mann, R. E., Smart, R. G., Stoduto, G., Adlaf, E. M., & Ialomiteanu, A. (2004). Alcohol consumption and problems among road rage victims and perpetrators. Journal of Studies on Alcohol, 65(2), 161-168.

Matthews, G., Dorn, L., Hoyes, T. W., Davies, D. R., Glendon, A. I., & Taylor, R. G. (1998). Driver stress and performance on a driving simulator. Human Factors: The Journal of the Human Factors and Ergonomics Society, 40(1), 136-149.

McCartt, A. T., Mayhew, D. R., Braitman, K. A., Ferguson, S. A., & Simpson, H. M. (2009). Effects of age and experience on young driver crashes: Review of recent literature

McDonald, C. C., Sommers, M. S., & Fargo, J. D. (2014). Risky driving, mental health, and health-compromising behaviours: Risk clustering in late adolescents and adults.0(1), 1.

McGinnis, J. M., Williams-Russo, P., & Knickman, J. R. (2002). The case for more active policy attention to health promotion. Health Affairs, 21(2), 78.

Miller, J. D., Campbell, W. K., Young, D. L., Lakey, C. E., Reidy, D. E., Zeichner, A., & Goodie, A. S. (2009). Examining the relations among narcissism, impulsivity, and self-defeating behaviors. *Journal of personality*, *77*(3), 761-794.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. Psychometrika, 50(2), 159-179.

Murphy, J. M., Horton, N. J., Monson, R. R., Laird, N. M., Sobol, A. M., & Leighton, A. H. (2014). Cigarette smoking in relation to depression: Historical trends from the Stirling County study. American Journal of Psychiatry, 160, 1663.

Musselwhite, C. (2006). Attitudes towards vehicle driving behaviour: Categorising and contextualising risk. Accident Analysis & Prevention, 38(2), 324-334.

Olfson, M., Kessler, R. C., Berglund, P. A., & Lin, E. (1998). Psychiatric disorder onset and first treatment contact in the United States and Ontario. American Journal of Psychiatry, 155(10), 1415-1422.

Petridou, E., & Moustaki, M. (2000). Human factors in the causation of road traffic crashes. European Journal of Epidemiology, 16(9), 819-826.

Petridou, E., Zavitsanos, X., Dessypris, N., Frangakis, C., Mandyla, M., Doxiadis, S., & Trichopoulos, D. (1997). Adolescents in high-risk trajectory: Clustering of risky behavior and the origins of socioeconomic health differentials. Preventive Medicine, 26(2), 215-219.

Sahai, V. S., Pitblado, J. R., Bota, G. W., & Rowe, B. H. (1998). Factors associated with seat belt use: An evaluation from the Ontario health survey. Canadian Journal of Public Health = Revue Canadienne De Sante Publique, 89(5), 320-324.

Schafer, J. L. (1999). Multiple imputation: A primer. Statistical Methods in Medical Research, 8(1), 3.

Scott, D., & Happell, B. (2011). The high prevalence of poor physical health and unhealthy lifestyle behaviours in individuals with severe mental illness. Mental Health Nursing, 32(9), 589.

Simpson, H., & Mayhew, D. (1992). Reducing the risks for new drivers: A graduated licensing system for British Columbia. Ottawa, Ontario: Traffic Injury Research Foundation.

Smart, R. G., Asbridge, M., Mann, R. E., & Adlaf, E. M. (2003). Psychiatric distress among road rage victims and perpetrators. Canadian Journal of Psychiatry, 48(10), 681-688.

Smart, R. G., & Mann, R. E. (2002b). Is road rage a serious traffic problem? Traffic Injury Prevention, 3(3), 183-189.

Sokal, R. R., & Sneath, P. H. A. (1963). Principles of numerical taxonomy. San Francisco, CA: W. H. Freeman & Co.

StataCorp. (2011). In College Station, TX: StataCorp LP. (Ed.), Stata statistical software: Release 12.

Statistics Canada. (2011). Canadian community health survey (CCHS) – annual component [microdata]. Ottawa, Ontario: Data Liberation Initiative [distributor].: Statistics Canada. Retrieved from http://www23.statcan.gc.ca/imdb-bmdi/document/3226_D7_T9_V8-eng.htm

Stoduto, G., Dill, P., Mann, R. E., Wells-Parker, E., Toneatto, T., & Shuggi, R. (2008a). Examining the link between drinking-driving and depressed mood. Journal of Studies on Alcohol and Drugs, 69(5), 777-780.

Transport Canada. (2005). Driver attitude to speeding and speed management: A quantitative and qualitative study – final report. (No. TP 14756 E). EKOS Research Associates Inc.

Transport Canada. (2013a). Canadian motor vehicle traffic collision statistics 2011 – TrafficCollisionStatisitcs_2011.pdf Retrieved from http://www.tc.gc.ca/media/documents/roadsafety/TrafficCollisionStatisitcs_2011.pdf

Transport Canada. (2014). Road safety in Canada. Retrieved from http://www.tc.gc.ca/eng/motorvehiclesafety/tp-tp15145-1201.htm

Transport Research International Documentation (TRID). (2013). Alcohol-crash problem in Canada: 2010. Retrieved from http://ccmta.ca/en/publications/resources-home/item/alcohol-crash-problem-in-Canada-2010

Ulleberg, P., & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. Safety Science, 41(5), 427-443.

Ulleberg, P. (2002). Personality subtypes of young drivers. Relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. Transportation Research Part F: Traffic Psychology and Behaviour, 4(4), 279-297.

Vanlaar, W., Robertson, R. D., & Marcoux, K. (2008). The road safety monitor 2007: Excessive speeding. Traffic Injury Research Foundation.

Vassallo, S., Smart, D., Sanson, A., Harrison, W., Harris, A., Cockfield, S., & McIntyre, A. (2007). Risky driving among young Australian drivers: Trends, precursors and correlates. Accident Analysis & Prevention, 39(3), 444-458.

Vingilis, E. (2007). Predictors of motor vehicle collision injuries among a nationally representative sample of Canadians. Traffic Injury Prevention, 8(4), 411.

Vingilis, E., & Wilk, P. (2010). Self-reported motor vehicle injury prevention strategies, risky driving behaviours, and subsequent motor vehicle injuries: Analysis of Canadian national population health survey. Canadian Public Policy, 36, 69-80.

Wang, P. S., Angermeyer, M., Borges, G., Bruffaerts, R., Tat Chiu, W., DE Girolamo, G., . . . Ustun, T. B. (2007). Delay and failure in treatment seeking after first onset of mental disorders in the world health organization's world mental health survey initiative. World Psychiatry: Official Journal of the World Psychiatric Association (WPA), 6(3), 177-185.

West, R., French, D., Kemp, R., & Elander, J. (1993). Direct observation of driving, self reports of driver behaviour, and accident involvement. Ergonomics, 36(5), 557-567.

White, M. P., Eiser, J. R., & Harris, P. R. (2004). Risk perceptions of mobile phone use while driving. *Risk Analysis, 24*, 323.

Wickens, C. M., Smart, R. G., & Mann, R. E. (2014). The impact of depression on driver performance. International Journal of Mental Health and Addiction, 12(4), 524.

Wickens, C. M., Mann, R. E., Stoduto, G., Ialomiteanu, A., Smart, R. G., & Rehm, J. (2013). The impact of probable anxiety and mood disorder on self-reported collisions: A population study. Journal of Affective Disorders, 145(2), 253-255.

Williams, J., Tregear, S., & Amana, A. (2011). Psychiatric disorders and driver safety: A systematic review. Paper presented at the Driving Assessment 2011: 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design,

Wilson, R. (1990). The relationship of seat belt non-use to personality, lifestyle and driving record. Health Education Research, 5(2), 175-185.

Zhao, N., Reimer, B., Mehler, B., D'Ambrosio, L. A., & Coughlin, J. F. (2013). Self-reported and observed risky driving behaviors among frequent and infrequent cell phone users. Accident Analysis & Prevention, 61, 71-77.

Zhou, R., Wu, C., Rau, P. –L. P., & Zhang, W. (2009). Young driving learners' intention to use a handheld or hands-free mobile phone when driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*(3), 208.

Ziedonis, D., Hitsman, B., Beckham, J. C., Zvolensky, M., Adler, L. E., Audrain-McGovern, J., & … Riley, W. T. (2008). Tobacco use and cessation in psychiatric disorders: National institute of mental health report. Nicotine & Tobacco Research, 10(12), 1691.

Chapter 5: Manuscript 2 (Pending Submission)

Profiling Risky Driving Behaviours Among Canadian Drivers – Associations With Mental Health

5.0 Abstract

Objective: The aim of this study was to explore associations between mental health factors and each of the five subgroups of driver risk behaviours revealed by *k*-means cluster analysis.

Methods: Data were based on a subsample drawn from the Driving and Safety module of the Canadian Community Health Survey (2011), a cross-sectional nationally representative sample of Canadians residents (n=47,356). *K*-means cluster analysis (performed a priori) revealed five subgroups of drivers: 1. Average Drivers (30.0%); 2. Cautious Drivers (33.3%); 3. Beltless Drivers (4.4%); 4. Egocentric Drivers (11.7%); and 5. Poly-risk Drivers (20.6%). Associations between four measures of mental health (diagnosis of a mood disorder, diagnosis of an anxiety disorder, stress, and negative mental health) and cluster membership were examined to assess whether mental health factors contribute to risky driving.

Analysis: Logistic regression analyses were performed on the four measures of mental health with cluster membership and demographic factors as independent variables.

Results: Associations between the mental health factors and cluster membership further differentiated the five clusters. Adjusted logistic regressions found the odds of positive responses to measures of poor mental health were significantly higher among clusters of risky drivers compared to clusters comprised of safer drivers. Significant associations were found between the subgroup of Beltless Drivers and all four measures of mental health. Aside from the Beltless Drivers, the odds of a diagnosis of a mood disorder (OR = 1.44, CI 1.23 – 1.70), higher levels of stress (OR = 1.40, CI 1.26 – 1.56), and negative mental health (OR = 2.42, CI 1.99 – 2.93) were higher among the riskiest subgroup of drivers (the Poly-risk Drivers) than among subgroups of safer drivers.

Conclusion: Mental health factors were linked to engagement in RDBs. Drivers with mood disorders, anxiety disorders, stress, and negative mental health were more likely to have engaged in multiple RDBs than drivers who did not report such factors. The findings of this study have important implications for safety of individuals with mental health disorders and the role of healthcare professionals, the workplace, and driver education in countermeasures to reduce engagement in RDBs and improve traffic safety.

5.1 Risky Driving Behaviours and Road Safety

Motor vehicle collisions (MVCs) are a major cause of disability and premature death. MVC research has widely recognized risky driving behaviours (RDBs) as an important contributor to MVC risk, particularly for those who speed, drive under the influence of alcohol (DUIA), and are seat belt non-compliant (Evans, 2004; Jonah, 1986; Petridou & Moustaki, 2000; Rajalin, 1994; Turner et al., 2004; Vingilis & Wilk, 2010). Improvements in traffic safety and the reduction of MVC-risk therefore greatly relies upon the prevention of RDBs (Petridou & Moustaki, 2000). Viewing crash risk in this way, as a function of RDBs, emphasizes the importance identifying the factors associated with RDBs. In particular, identifying whether individual factors are associated with specific RDBs or if they are common to multiple interrelated driving behaviours is essential to accurate profiling of drivers who engage in RDBs, identifying high-risk drivers, and implementation of prevention strategies (Petridou & Moustaki, 2000).

Studies on risk propensity show that risk-taking behaviours tend to coexist and are often apparent across multiple health-compromising behaviours (Anderson & Mellor, 2008; Dohmen et al., 2011; Jessor, 1987; McDonald et al., 2014a; Petridou et al., 1997). The evidence of risk taking across different health behaviours suggests that multiple risk-taking behaviours are likely to exist across driving behaviours (drivers who take risks may engage in multiple RDBs). Yet, little research has explored whether RDBs coexist among the same drivers using a comprehensive range of RDBs, nor has it explored the factors associated with such patterns of driving behaviour. Furthermore, although research has indicated links between poor mental health and many risk-taking behaviours such as smoking, unsafe sexual behaviours, and poor diet and physical activity, very little research has empirically explored the association between mental health factors and engagement in RDBs (Murphy et al., 2014; Scott & Happell, 2011; Wickens et al., 2014; Ziedonis et al., 2008). Determining whether RDB coexist among risky drivers as patterns of RDBs and clarifying the factors, including mental health factors, associated with risky drivers is essential to accurately identifying high-risk drivers. Research of RDBs has not yet confirmed whether the 'homogeneity hypothesis' exists in the context of RDBs, whereby drivers who engage in a specific RDB are not distinct from other risky drivers, but belong to a larger profile of drivers who engage in multiple RDBs.

The objectives of this study was to explore the association between self-reported RDBs of drivers in each of the five subgroups (clusters) revealed by a *k*-means cluster analysis and four self-reported mental health-related measures – diagnosis of a mood disorder, diagnosis of an anxiety disorder, stress, and negative mental health – using logistic regression, controlling for the effects of sociodemographic measures, in a large population of Canadian drivers. It was hypothesized that associations exist between cluster membership and drivers' self-reported measures of mental health, whereby clusters composed of riskier drivers would associate with higher reported rates of poor mental health.

5.2 Literature Review

Research shows that the two most common psychiatric disorders affecting the population, mood and anxiety disorders, in addition to stress and negative mental health in general, may have detrimental effects on driving behaviour, skills, and risk of MVC (Hu, Xie, & Li, 2013; Kessler & Wang, 2008; Mann et al., 2010; Rowden, Matthews, Watson, & Biggs, 2011; Stoduto et al., 2008; Wickens et al., 2013; Williams et al., 2011).

The association between mental health-related factors and RDBs, however, remains unclear. The number of studies exploring mental health factors and their contribution to RDBs have either failed to consider confounding variables, have had limited sample sizes, or have varied in their definitions of measures of mental health (Williams et al., 2011). Furthermore, research to date has yet to explore the association between mental health factors and patterns of RDBs in a large population of drivers of all ages.

Mental health-related factors that have received some empirical attention are depression and measures of psychiatric distress and anxiety (Butters et al., 2006; Fong, Frost, & Stansfeld, 2001; Hu et al., 2013; Smart et al., 2003; Vaughn et al., 2011). A cluster analysis of 2,610 drivers from Ontario by Smart, Asbridge, Mann, and Adlaf (2003), demonstrated that drivers who reported severe road rage involvement scored higher on the GHQ measure of psychiatric distress than drivers with little or no involvement in road rage. Other research has suggested links between depressed mood and DUIA (Stoduto et al., 2008; Wickens et al., 2014). In a telephone survey of 3,979 drivers in Ontario, Stoduto, et al. (2008) used logistic regression to examine the association between depressed mood (GHQ-12) and self-reported DUIA (driving after two or more alcoholic drinks in the previous hour). After controlling for measures of alcohol use, driving exposure, and demographic factors, results showed that the odds of DUIA increased significantly as scores measuring depressed mood increased (OR = 1.08, 95% CI: 1.03 – 1.12).

Studies investigating the contribution of anxiety to RDBs is mixed (Dula, Adams, Miesner, & Leonard, 2010; Ge et al., 2014; Ulleberg, 2002). Dula, Adams, Miesner, and Leonard (2010) demonstrated that, independent of sex, a higher anxiety level was a factor associated with most RDBs. In this study, 1,121 participants were assigned to one of three groups, low, medium, and high anxiety according to their scores from the Beck Anxiety Inventory (BAI) (Beck, Epstein, Brown, & Steer, 1988). MANOVA tests revealed significant differences between the three groups for the majority of measures of risky driving (the Dula Dangerous Driving Index and the Propensity for Angry Driving Scale). Results showed significant differences in the number of DUIA episodes in the previous year between the low ($M = 0.95$, $SD = 3.99$) and medium anxiety groups ($M = 1.29$, $SD = 5.84$) and the high anxiety group ($M = 3.21$, $SD = 14.63$, $F = 9.06$, $p<0.00$) and the number of seat belt citations in the previous five years between the low anxiety group ($M = 0.52$ $SD = 1.10$) and high anxiety group ($M = 0.73$, $SD = 1.46$, $F = 3.07$, $p = 0.05$).

Other studies have suggested a non-linear association between anxiety and RDBs. In a study measuring the impact of personality on RDB and MVC risk, Oltedal and Rundmo (2006) found a weak association between anxiety and RDBs, and proposed that the association between anxiety and risky driving may be U-shaped, with a tendency for both high and low anxiety drivers to exhibit RDBs. The authors suggested that the non-linear relationship may result from high anxiety drivers experiencing excessive tension and cognitive interference and low anxiety drivers' lack of concern about driving safety and over confidence. Previous research had exemplified this "U" shaped association; a cluster analysis of personality measures by Ulleberg, 2002, demonstrated that measures of sensation seeking, aggression, anxiety, altruism, driving anger, and normlessness clustered into six distinct groups that differed according to measures of attitudes towards traffic safety, self-evaluation of driving skills, risk perception, and accident involvement.

Two of the six sub-groups of drivers were identified as high-risk clusters of drivers, one with low anxiety and one with high anxiety. The first high-risk cluster contained mostly male drivers with low anxiety and altruism and high levels of sensation seeking, irresponsibility, and driving aggression, while the other cluster was comprised of predominantly female drivers with high anxiety, sensation seeking, aggression, and driving anger.

The association between driver stress and risky driving is complex, with personality and cognitive factors and the ecological relationship between person and environment playing important roles (Matthews, 2002). Despite the complex nature of stress, research has demonstrated that various forms of stress such as driving-related stress, work stress, and global stress, negatively impact driving behaviour (Rowden et al., 2011; Westerman & Haigney, 2000). Although it is difficult to determine whether stress outside the driving environment impacts driving behaviours independently of driving-related stress, some research such as that by Rowden, Matthews, Watson, and Biggs (2011) has shown such an association. In their study, Rowden, et al. (2011) examined the impact of stress on driving behaviour and road safety of 247 Australian drivers. Confirmatory factor analysis and structural equation modelling based on participants' responses to self-reported measures of subjective work-related stress, daily hassles, and aspects of general mental health [from the Driver Behavior Questionnaire (DBQ) and Driver Stress Inventory (DSI)], in addition to demographic measures, showed associations between driver stress and risky driving (DBQ errors, lapses, and violations). In addition, results indicated that negative affect, risk taking, and extraneous influences (daily hassles, work-related stress, and general mental health) were predictive of risky driving. Importantly, results demonstrated that daily hassles independently impacted driver behaviour, specifically driver lapses (standardized path coefficient = .21) and driver violations (standardized path coefficient = .13).

A better understanding of the contribution of mental health factors such as depression, anxiety, stress, and negative mental health to RDBs may have important implications for traffic safety and the wellbeing of individuals reporting such factors. This study explored the associations between of four mental health-related variables, diagnosis of a mood disorder, diagnosis of an anxiety disorder, stress, and negative mental health and five subgroups of drivers generated by a *k*-means cluster analysis of six RDBs performed a priori. Exploration of the contribution of mental health factors to a broad range of self-reported RDBs among drivers of all ages will help to further demarcate the factors associated with particular forms or patterns of risky driving.

5.3 Methodology

5.3.1 Data

This study utilized a subsample of data from the Canadian Community Health Survey – Annual component, 2011 (CCHS) that looked at risky driving (Statistics Canada, 2011). The CCHS is a cross-sectional nationally representative sample of Canadians residents aged 12 years and older performed by Statistics Canada between January 2009 and December 2010. Data pertaining to health status, health care utilization and determinants of health were collected via face to face and telephone interviews and achieved an overall response rate of 72.3%. Residents of Indian Reserves, Canadian Forces Bases, institutional accommodations and some individuals living in extremely remote areas were excluded from the sampling frame.

The subsample of data was drawn from the CCHS optional module Driving and Safety, which included questions related to risky driving (Statistics Canada, 2011). This module was administered to a total of 59,163 residents living in Newfoundland (3,768), Ontario (42,495), Alberta (11,618), and the Yukon Territory (1,282), in an effort to understand safe driving practices and reduce the risk of MVC and associated morbidity and mortality. For this study, subjects included those 16 years and older who had driven a motor vehicle (including a car, truck, or van) in the previous year.

5.3.2 Independent Measures

Independent measures included sociodemographic variables of interest drawn from the CCHS core component in addition to cluster membership generated by a cluster analysis (performed a priori) of six questions from the Driving and Safety optional module of the CCHS.

5.3.2.1 Sociodemographic Measures

Age, sex, geographic location, race, marital status, education, income, employment status, and immigrant status have been linked to engagement in various RDBs and mental health factors (Government of Canada, 2006; Transport Canada, 2013b). Sex and age were two important measures included as independent measures, as young age and male sex are linked with most RDBs, while older age and female sex are associated with poorer mental health. Self-reported age in years was reported according to 16 categories of two and four year increments between the ages of 12 and 80 years or older. The variable age was recoded into the following categories: (a) less then 20 years; (b) 21 – 25 years; (c) 26 – 40 years; (d) 41-55 years; (e) 56 – 70 years; and (f) 71 years and older, with less than 20 years as the referent category. Dichotomous variables included *sex*, coded as (0) female (referent) and (1) male; *geographic location* coded as (0) rural (referent) and (1) urban; *immigrant status*, coded (0) non-immigrant (referent) and (1) immigrant; and *employment*, coded (0) employed and (1) unemployed. *Marital status* was captured by the CCHS (2011) as married, common-law, widow/separated/divorced, or single, never married. To meet the minimum required cell sizes in accordance to the CCHS mandate to ensure confidentiality, the variable marital was dichotomized as (0) married or living common-law (referent) and (1) not married. Two variables served as proxies for socioeconomic status (SES). Missing categories for *race*, *education*, and *income* were included due to a large percentage of missing data for these variables. Cultural or racial origin (*race*) was coded (1) White (referent); (2) non-White; (3) missing; *education* was coded (1) less than secondary education; (2) secondary education; (3) some post-secondary education (referent); (4) post-secondary graduate; and (5) missing; and annual household income (*income*) was coded (1) $0-19,999; (2) $20,000 - $39,999; (3) $40,000 - $59,999; (4) $60,000 - $79,999; and (5) $80,000 or more (referent); (6) missing.

5.3.2.2 Cluster Membership

Cluster analysis was performed a priori based on six questions from the Driving and Safety optional module of the CCHS. Participants who reported driving a motor vehicle in the previous 12 months responded to questions that assessed seat belt non-compliance, cell phone-distracted driving, fatigued driving, speeding, aggressive driving, and DUIA. The response options for the regularity of seat belt use included (a) always; (b) sometimes; (c) rarely; and (d) never. Response categories for the frequency of cell phone use while driving and the frequency of fatigued driving each included (a) often; (b) sometimes; (c) rarely; and (d) never. Speeding and aggressive driving were measured similarly and included the response categories (a) much faster/much more aggressively; (b) a little faster/a little more aggressively; (c) about the same; (d) a little slower/a little less aggressively; and (e) much slower/aggressively than other drivers. DUIA was defined as having two or more drinks within one hour prior to driving a motor vehicle and was measured according to the response categories (a) yes and (b) no.

A two-stage cluster analysis (Ward's hierarchical agglomerative and *k*-means) revealed a clear pattern in RDBs. The cluster solution contained five significantly different clusters of risky drivers (Table 8). A series of one-way ANOVA tests in addition to post-hoc Bonferroni tests confirmed significant differences between clusters. The five-cluster solution showed two patterns of very risky driving: 1. the Poly-risk Drivers (n = 9,745) who often drove while fatigued, drove much faster and much more aggressively than other drivers, and sometimes engaged DUIA; and 2. the Egocentric Drivers (n = 5,558) who frequently engaged in cell phone-distracted driving, sometimes drove while fatigued, drove a little faster and a little more aggressively than other drivers, and reported DUIA. A third subgroup of moderately risky drivers, 3. the Average Drivers (n = 14,179), were seat belt compliant, reported driving at speeds and levels of aggression comparable to other drivers, and did not report DUIA, but occasionally engaged in cell phone-distracted driving and fatigued driving. The fourth subgroup of drivers 4. The Cautious Drivers (n = 15,791) represented the safest drivers who were the refrained from all forms of RDBs. The Cautious Drivers were seat belt compliant, drove much slower and much less aggressively than other drivers, never engaged in fatigued driving, or DUIA, but did report rare cell phone use while driving. The fifth and smallest cluster, 5. the Beltless Drivers (n = 2,083), did not engage in any of the RDBs aside from seat belt non-compliance.

The five-cluster solution differed not only according to RDBs, but also in sociodemographic, health, and other risk taking-related variables. Results showed significant differences between clusters for the mental health-related variables (Table 15).

5.3.3 Dependent Measures

Considering recent research linking psychiatric factors such as depression, anxiety, and stress to the engagement in RDBs, and given the associations between the mental health factors and the five subgroups generated by cluster analysis (a priori), four mental health-related variables were used as dependent measures in the analysis (Dula et al., 2010; Freeman, Maxwell, & Davey, 2011; Hu et al., 2013; Malta, Blanchard, & Freidenberg, 2005; Wickens et al., 2014; Wickens et al., 2013). These included: 1. *diagnosis of a mood disorder*, "Remember, we are interested in conditions diagnosed by a health professional. Do you have a mood disorder such as depression, bipolar disorder, mania or dysthymia?"; measured as *mood* (no mood vs. mood); 2. *diagnosis of an anxiety disorder*, "Do you have an anxiety disorder such as a phobia, obsessive-compulsive disorder or a panic disorder?"; measured as *anxiety* (no anxiety vs. anxiety); 3. *perceived life stress*, "Thinking about the amount of stress in your life, would you say that most days are: (not at all stressful, a bit stressful, quite a bit stressful, or extremely stressful)?"; measured as *stress* (lower levels of stress vs. higher levels of stress); and 4. *perceived mental health*, "In general, would you say your mental health is: (excellent, very good,

good, fair, or poor?"; measured as *mental health* (positive mental health vs. negative mental health) (Statistics Canada, 2011).

5.4 Analysis

Probability sampling weights were used in all analyses to produce population estimates at the health region level for Canadian drivers 16 years and older. All analyses were conducted using STATA version 12.0 (StataCorp., 2011). STATA's robust option was used to adjust standard errors for survey design effects resulting from the CCHS complex sampling design.

5.4.1 Multivariate Analyses

Four separate unadjusted logistic regression analyses (CI 95%) were performed, regressing each of the mental health-related variables (mood, anxiety, higher levels of stress, and negative mental health) on risky driving clusters with the Average Drivers as the referent. Logistic regressions (CI 95%) were then repeated, adjusting for sociodemographic variables. Missing variable categories were included for race, education, and income to maintain the overall study population, due to STATA's use of listwise deletion of missing data.

5.5 Results

The logistic regression results demonstrated significant associations between cluster membership and each of the four mental health-related variables. Table 16 provides the OR, beta coefficient (*B*), and 95% confidence intervals (CI) from the four unadjusted logistic regressions of each mental health variable on cluster membership. Results showed the odds of a mood disorder diagnosis were higher among the Beltless Drivers (OR 1.77, 95% CI 1.42 – 2.21) and Poly-risk Drivers (OR 1.37, CI 1.16 – 1.60) compared to the Average Drivers, and there was a 17% decrease in the odds of a mood disorder diagnosis among the Egocentric Drivers (OR 0.83, CI 0.68 – 1.00). The only cluster with a significant association with a diagnosis of an anxiety disorder was the cluster of Beltless Drivers (OR = 1.81, CI 1.39 – 2.34). The Egocentric Drivers (OR = 1.58, CI 1.41 – 1.78) and Poly-risk Drivers (OR = 1.59, CI 1.43 – 1.77) were both associated with higher levels of stress compared to the Average Drivers, and there was a 29% decrease in the odds of higher levels of stress among the Cautious Drivers (OR = 0.71, CI 0.66 – 0.77). Finally, the odds ratios for negative mental health were highest among the Beltless Drivers (OR = 29.14, CI 23.90 – 35.51) followed by the Poly-risk Drivers (OR = 2.10 CI 1.74 – 2.54).

Tables 17 – 20 summarize the results of the logistic regressions of the four mental health-related variables on driving clusters while controlling for sociodemographic factors. The adjusted results maintained significance for most of the cluster variables, independent of the sociodemographic factors included in the models.

Table 17 presents the results of the adjusted logistic regressions for mood disorders. Significant associations were maintained between mood disorder and the Beltless Drivers (OR = 1.62, 95% CI 1.26 – 2.09) and Poly-risk Drivers (OR = 1.44, CI 1.23 – 1.70). The OR for Egocentric Drivers was no longer significant in the adjusted model (OR = 0.99, CI 0.81 – 1.22).

Table 18 summarizes the adjusted logistic regressions for diagnosis of an anxiety disorder on driver clusters and sociodemographic variables. Results show the Beltless Drivers had the highest odds of reporting an anxiety disorder (OR = 1.60, 95% CI 1.19 – 2.16); there were no significant associations observed for other driver clusters.

The results of the adjusted logistic regression of higher levels of stress on cluster membership are summarized in Table 19. Relative to Average Drivers, the Poly-risk Drivers (OR = 1.40, 95% CI 1.26 – 1.56), Egocentric Drivers (OR = 1.36, CI 1.20 – 1.53), and Beltless Drivers (OR 1.32, CI 1.10 – 1.58) were all associated with higher levels of stress. Compared to the Average Drivers, the Cautious Drivers showed a 16% decrease in the odds of higher levels of stress.

Finally, Table 20 summarizes the adjusted logistic regression results of fair or poor mental health on cluster membership. Relative to Average Drivers, the Beltless Drivers (OR = 39.51, 95% CI 30.73 – 50.80) and Poly-risk Drivers (OR = 2.42, CI 1.99 – 2.93) reported significantly higher odds of fair/poor mental health. The ORs associated with the Cautious Drivers in the unadjusted results were no longer significant in the adjusted model (OR = 0.98, CI 0.81 – 1.18).

5.6 Discussion

This study explored the associations between each of four mental health variables and engagement in RDBs, using a *k*-means cluster solution of six RDBs to generate five heterogeneous subgroups of drivers. It was hypothesized that associations would be found between positive responses to the measures of mental health and membership in the subgroups of risky drivers (relative to the Average Drivers subgroup). As expected, negative mental health as well as a diagnosis of a mood disorder were both associated with membership in the subgroups of Poly-risk Drivers and the Beltless Drivers. Diagnosis of an anxiety disorder was significantly associated with membership in the cluster of Beltless Drivers. Higher levels of stress were significantly associated with membership in all subgroups of risky drivers, with membership in the Cautious Drivers subgroup demonstrating a protective effect. These links between risky driving and mental health factors are consistent with findings of previous studies (Matthews et al., 1998; Rowden et al., 2011; Wickens et al., 2014; Wickens et al., 2013).

5.6.1 Mental Health by Driver Profile

5.6.1.1 The Poly-risk Drivers

The measures of mental health associated with the riskiest subgroup of drivers, the Poly-risk Drivers, highlighted the important role of mental health factors in RDBs. Being in this this subgroup were associated with a diagnosis of a mood disorder, higher levels of stress, and negative mental health, which confirmed and extended previous investigations linking drivers with psychiatric disorders, higher levels of stress, or negative mental health to dangerous driving behaviours, particularly aggressive driving and speeding (Freeman et al., 2011; Malta et al., 2005; McDonald, Sommers, & Fargo, 2014b; Smart et al., 2003; Wickens et al., 2014; Williams et al., 2011; Yu, Evans, & Perfetti, 2004). Studies that have explored associations of mental health factors with driving behaviour have demonstrated a negative impact, such as deficits in driving skills that result in an increase in driving errors and aggressive responses during frustrating driving situations (Ramaekers, Ansseau, Muntjewerff, Sweens, & O'Hanlon, 1997; Stephens & Groeger, 2009; Yu et al., 2004). For example, in a review of the literature of the impact of depression on driving behaviour, Wickens, et al. (2014) found that depressed drivers have difficulty with reaction time, changes in speeds when following another vehicle, divided attention, and weaving within the lane.

5.6.1.2 The Egocentric Drivers

Previous research has shown that higher levels of stress negatively contribute to multiple forms of risky driving (Rowden et al., 2011). The association between higher levels of stress and the second riskiest subgroup of drivers, the Egocentric Drivers supported such findings and were of particular interest, as this subgroup of drivers was not associated with any other measures of mental health. The Egocentric Drivers consisted predominantly of young, male, employed, and unmarried drivers, who live in urban areas. These factors suggest that such drivers may live fast-paced and busy lifestyles, perhaps indicative of individuals working in high-stress business careers in urban centres, resulting in a higher susceptibility to stress-related symptoms that affect driving behaviours. As Rowden, et al. (2011) describes, individuals experiencing stress resulting from work overload, inter-personal conflicts, and role ambiguity may exhibit such symptoms as fatigue, nervousness, alcohol abuse, or drug abuse, which are known to detrimentally influence driving abilities.

5.6.1.3 The Beltless Drivers

The findings for the Beltless Drivers were particularly interesting, as these drivers refrained from all forms of RDBs, aside from seat belt non-compliance; yet, membership to this unique subset of drivers was associated with positive responses to all four measures of mental health. Furthermore, the subgroup of Beltless Drivers was the only driving cluster associated with a diagnosis of an anxiety disorder and had almost 30 times higher odds than the Average Drivers of reporting negative mental health in the unadjusted analysis. These findings may be explained by the high mean age of the drivers in this cluster. Multiple diagnoses of psychiatric disorders, higher levels of stress, negative mental health, and a range of other physical health issues are more common among older than younger populations. Furthermore, the prevalence of the use of mental health services among Canadians 20 years and older is highest for individuals 80 years and older (24.1%) (Government of Canada, 2006).

Another explanation for the findings associated with the Beltless Drivers may be marital dissolution or widowhood. Previous research has suggested that stressful life events such as separation, divorce, or death of a spouse are factors associated with an increased risk of psychiatric distress and disorders such as mood disorders and anxiety (Bhatti et al., 2008; Charlton et al., 2003; Charlton et al., 2006; Mazure, 1998; Sbarra, Emery, Beam, & Ocker, 2014; Useche, Serge, & Alonso, 2015). For example, in Canada, individuals with a spouse or partner are more likely to be mentally healthy than individuals who are separated, divorced, widowed, (OR 0.8 CI 95% 0.7 – 1.0), or never married (OR 0.7 CI 95% 0.6 – 0.8) (Gilmore, 2014). Given the Beltless Drivers' relatively high mean age, marital dissolution or widowhood may have contributed to the high proportion of drivers who reported mental health issues.

Retirement status is another variable unaccounted for by this study. Stable income and employment are known protective factors for mental health, two factors that are not characteristic of the Beltless Drivers. In fact, the Beltless Drivers were least likely to have annual household incomes of more than $80,000 and least likely to be employed, suggesting that this subgroup of drivers is no longer professionally active, perhaps helping to explain why these drivers are the most likely of all clusters to report poor mental health (Government of Canada, 2006). In addition, research shows that retired drivers tend to engage in fewer RDBs than employed drivers. According to Bhatti, et al. (2008) retired drivers are more likely to discontinue fatigued driving (OR =

2.12, $p = 0.001$) and cell phone-distracted driving (OR = 1.74, $p = 0.006$) than employed drivers. This may explain why the Beltless Drivers do not report any other RDBs, aside from seat belt non-compliance. Whether retirement is a factor associated with the poor mental health reported by the Beltless Drivers, or with the lack of RDBs, or both, this study would have benefited from the inclusion of a measure of retirement status.

5.6.1.4 The Cautious Drivers

In contrast to the subgroups of drivers who exhibit risky driving, the safest subgroup of drivers, the Cautious Drivers, refrained from RDBs; rather, they reported driving slower and less aggressively than other drivers. In addition, this subgroup of drivers was not associated with responses indicating mental health problems. In fact, the Cautious Drivers were significantly less likely to report higher stress than the Average Drivers.

5.6.1.5 The Average Drivers

This study explored measures of mental health using the largest subgroup of drivers, the Average Drivers, as the referent category. It appears that the Average Drivers had significantly higher levels of stress, and slightly (not statistically significant) greater odds of diagnosis of a mood disorder, than drivers in the subgroup of Cautious Drivers. Similar to the riskier subgroups of drivers, the Average Drivers occasionally engaged in cell phone-distracted driving and fatigued driving.

5.7 Limitations

There are important limitations to consider in the interpretation of the findings of this research. First, although the findings of this study are of substantial interest, cluster analysis is a descriptive non-inferential procedure useful for exploratory data analyses of large data sets, which restricts confirmation of the associations between RDBs and the measures of mental health. Second, the results of the study were based upon self-reported data. Although self-report RDBs are shown to reliably reflect actual driving behaviour, the possibility of response bias may limit results, particularly for driving behaviours deemed socially unacceptable and illegal such as DUIA (Lajunen & Summala, 2003; West, French, Kemp, & Elander, 1993). Third, variables unaccounted for may have influenced results. The inclusion of measures related to illegal drug use, retirement status, and measure of marital status inclusive of categories reflecting divorce, separation, and widowhood may benefit future research and generate a more accurate assessment of the association between RDBs and mental health factors. Fourth, this study did not consider the impact of psychiatric medications on driving behaviours. The medications prescribed to treat many psychiatric disorders have been associated with cognitive and psychomotor disruptions that may negatively affect driving skills (Hindmarch, Alford, Barwell, & Kerr, 1992; Taylor, Deane, & Podd, 2008). Finally, the associations found between patterns of RDBs and the measures of mental health may have been conservative and underestimated the contribution of mental health to risky driving. Assigning the subgroup of Cautious Drivers as the referent category as opposed to the Average Drivers may have resulted in a more accurate assessment of the associations between patterns of risky driving and the four measures of mental health. Future cluster analyses using similar methodology should carefully consider the optimal referent category (largest or safest subgroup of drivers) in their analyses.

5.8 Conclusion

A five-cluster solution revealed three patterns of RDBs, a small subset of seat belt non-compliant drivers, and a subgroup of safe drivers. The present study explored the associations between patterns of RDBs and four mental health factors, using multiple logistic regressions while controlling for a variety of sociodemographic measures. The findings demonstrate that cluster membership is associated with measures of mental health. The associations between each subgroup of drivers and the measures of mental health further differentiate the five clusters of drivers and provide external validation to the five-cluster solution.

Broadly, this study suggests that risky drivers are more likely to report a diagnosis of a mood disorder, higher stress, and negative mental health than safer drivers and highlights the importance of understanding the link between mental health and risky driving. Inclusion of measures of mental health in future MVC research will not only clarify the link between mental health and RDBs and improve traffic safety, but will also have important implications for the safety of individuals with mental health disorders. In fact, health care settings and professionals may play an important role in more accurately identifying high-risk drivers, particularly those with mental health disorders, and targeting education of the risks associated with RDBs. The workplace may also play an important role in increasing awareness of the impact of stress and mental health on driving behaviours and improving the traffic safety culture associated with cell phone-distracted driving and DUIA. Finally, driver education programs and provincial departments of motor vehicles may benefit from the inclusion of an educational component related to the impact of mental health on driving behaviour and safety in addition to considering the importance of seat belt compliance when educating, licensing, or re-licensing older drivers.

Table 15
K-means five-cluster solution

RDB	N (%)	Average Drivers[a] 14,179 (30.0)	Cautious Drivers[b] 15,791 (33.3)	Beltless Drivers[c] 2,083 (4.4)	Egocentric Drivers[d] 5,558 (11.7)	Poly-risk Drivers[e] 9,745 (20.6)	F Score[f]
Seat belt◆							
M^g (SD^h)	3.81 (0.66)	3.95 (0.24)	3.97 (0.16)	1.04 (0.19)	3.84 (0.48)	3.92 (0.36)	46,288*
CI^i (99%)	(3.80, 3.82)	(3.95, 3.95)	(3.97, 3.98)	(1.03, 1.05)	(3.83, 3.85)	(3.91, 3.93)	
Cell Phone							
M (SD)	3.34 (0.95)	3.66 (0.47)	3.66 (0.72)	3.94 (0.36)	1.59 (0.49)	3.58 (0.59)	19,807*
CI (99%)	(3.33, 3.35)	(3.65, 3.67)	(3.64, 3.67)	(3.92, 4.00)	(1.58, 1.60)	(3.58, 3.60)	
Fatigue◆							
M (SD)	2.94 (0.94)	3.00 (0.04)	4.00 (0.06)	3.98 (0.14)	2.32 (0.67)	1.72 (0.45)	78,490*
CI (99%)	(2.93, 2.95)	(3.00, 3.00)	(4.00, 4.00)	(3.97, 3.99)	(2.30, 2.33)	(1.71, 1.73)	
Speeding◆							
M (SD)	2.99 (1.00)	3.00(0.04)	4.00 (0.09)	4.88 (0.38)	2.32 (0.6)	1.72 (0.45)	85,333*
CI (99%)	(2.98, 3.00)	(3.00, 3.00)	(4.00, 4.00)	(4.87, 4.90)	(2.30, 2.33)	(1.71, 1.73)	
Aggressive Driving◆							
M (SD)	2.99 (1.01)	3.00 (0.08)	4.02 (0.13)	4.88 (0.37)	2.32 (0.67)	1.72 (0.46)	81,991*
CI (99%)	(2.98, 3.00)	(3.00, 3.00)	(4.02, 4.02)	(4.87, 4.90)	(2.30, 2.34)	(1.71, 1.73)	
DUIA							
M (SD)	0.05 (0.22)	0.04 (0.02)	0.02 (0.14)	0.01 (0.01)	0.13 (0.34)	0.07 (0.25)	328*
CI (99%)	(0.05, 0.05)	(0.04, 0.04)	(0.02, 0.02)	(0.01, 0.01)	(0.12, 0.14)	(0.06, 0.07)	

[a] Cluster 1, [b] Cluster 2, [c] Cluster 3, [d] Cluster 4, [e] Cluster 5, [f] F score generated by ANOVA tests evaluated between group differences in risky driving behaviours in the *K*-means cluster solution, df (4, 47351), [g] mean, [h] standard deviation, [i] confidence interval, *$p<.001$◆
◆Post hoc Bonferroni tests [df (4, 47351), p<.001] confirmed pairwise comparisons were significantly different between all five clusters

Table 16
ANOVA results comparing mental health variables in the *k*-mean five-cluster solution

Variable (Code range)	*N*	Average Drivers[a]	Cautious Drivers[b]	Beltless Drivers[c]	Egocentric Drivers[d]	Poly-Risk Drivers[e]	*F* Score[f]
n	47,356	14,179	15,791	2,083	5,558	9,745	
Mental Health Status (0-1)							
M (*SD*)[g]	0.07 (0.25)	0.04 (0.19)	0.05 (0.20)	0.55 (0.49)	0.03 (0.21)	0.06 (0.25)	2,330.96*
CI (99%)	(0.07, 0.07)	(0.04, 0.04)	(0.04, 0.05)	(0.52, 0.58)	(0.03, 0.04)	(0.06, 0.07)	
Stress (0-1)							
M (*SD*)	0.67 (0.47)	0.66 (0.46)	0.58 (0.46)	0.64 (0.48)	0.75 (0.49)	0.77 (0.44)	305.40*
CI (99%)	(0.66, 0.68)	(0.65, 0.67)	(0.57, 0.59)	(0.62, 0.67)	(0.74, 0.77)	(0.76, 0.78)	
Mood Disorder (0-1)							
M (*SD*)	0.07 (0.25)	0.06 (0.24)	0.06 (0.22)	0.11 (0.30)	0.05 (0.25)	0.08 (0.28)	30.90*
CI (99%)	(0.07, 0.07)	(0.06, 0.07)	(0.05, 0.06)	(0.09, 0.12)	(0.04, 0.06)	(0.07, 0.09)	
Anxiety Disorder (0-1)							
M (*SD*)	0.05 (0.21)	0.04 (0.20)	0.04 (0.19)	0.07 (0.26)	0.04 (0.23)	0.05 (0.23)	14.16*
CI (99%)	(0.05, 0.05)	(0.04, 0.05)	(0.04, 0.05)	(0.06, 0.09)	(0.04, 0.05)	(0.05, 0.06)	

[a] Cluster 1, [b] Cluster 2, [c] Cluster 3, [d] Cluster 4, [e] Cluster 5, [f] *df* (4, 47351), *p*<.001, [g] mean (standard deviation), [h] confidence interval.
**p<0.01;* Bonferroni post-hoc tests demonstrated significant differences in the mean scores for each mental health variable for all five clusters [*df* (4, 47351), *p*<.001].

Table 17
Unadjusted logistic regression of diagnosis of a mood disorder, diagnosis of an anxiety disorder, higher levels of stress, and negative mental health on driving clusters (p < .05)

Independent variables	Mood		Anxiety		Higher levels of stress		Negative Mental Health	
K-means cluster	*OR*[a] (*B*)[b]	*CI*[c] (95%)	*OR* (*B*)	*CI* (95%)	*OR* (*B*)	*CI* (95%)	*OR* (*B*)	*CI* (95%)
Average drivers	Ref*		Ref		Ref		Ref	
Cautious drivers	0.94 (-0.07)	(0.80, 1.09)	0.99 (-0.01)	(0.83, 1.19)	0.71 (-0.34)	(0.66, 0.77)	1.20 (0.18)	(1.00, 1.44)
Beltless drivers	1.77 (0.57)	(1.42, 2.21)	1.81 (0.59)	(1.39, 2.34)	0.92 (-0.08)	(0.79, 1.08)	29.14 (3.37)	(23.90, 35.51)
Egocentric drivers	0.83 (-0.19)	(0.68, 1.00)	0.97 (-0.03)	(0.76, 1.25)	1.58 (0.46)	(1.41, 1.78)	0.85 (-0.16)	(0.68, 1.07)
Poly-risk drivers	1.37 (0.31)	(1.16, 1.60)	1.05 (0.05)	(0.87, 1.27)	1.59 (0.46)	(1.43, 1.77)	2.10 (0.74)	(1.74, 2.54)
Constant	0.07 (-2.71)	(0.06, 0.07)	0.04 (-3.11)	(0.04, 0.05)	1.94 (0.66)	(1.83, 2.06)	0.04 (-3.18)	(0.04, 0.05)

[a] Odds ratio; [b] Beta coefficient; [c] Confidence interval
* Referent category

Table 18
Logistic regression of diagnosis of a mood disorder on driving clusters and sociodemographic characteristics ($p < .05$)

Independent variables	Cluster	OR^a $(B)^b$	CI^c (95%)
K-means cluster	Average Drivers	Ref*	
	Cautious Drivers	0.91 (-0.10)	(0.78, 1.07)
	Beltless Drivers	1.62 (0.49)	(1.26, 2.09)
	Egocentric Drivers	0.99 (-0.01)	(0.81, 1.22)
	Poly-risk Drivers	1.44 (0.37)	(1.23, 1.70)
Age	<=20	Ref	
	21-25	1.52 (0.42)	(1.08, 2.15)
	26-40	2.37 (0.86)	(1.76, 3.20)
	41-55	2.90 (1.07)	(2.16, 3.90)
	56-70	2.08 (0.73)	(1.54, 2.81)
	71+	0.84 (-0.18)	(0.61, 1.15)
Sex	Male	0.61 (-0.50)	(0.54, 0.68)
Location	Urban	1.09 (0.09)	(0.96, 1.25)
Race	White	Ref	
	Non-White	0.44 (-0.83)	(0.33, 0.57)
	Missing	1.15 (0.14)	(0.89, 1.49)
Marital status	Unmarried	0.73 (-0.31)	(0.64, 0.83)
Education	Less than secondary	0.92 (-0.08)	(0.74, 1.16)
	Secondary graduate	0.85 (0.16)	(0.70, 1.04)
	Some post-secondary	1.14 (0.13)	(0.88, 1.46)
	Post-secondary graduate	Ref	
	Missing	0.91 (-0.10)	(0.68, 1.21)
Income	$0-$19,999	2.40 (0.88)	(1.89, 3.05)
	$20,000-$39,999	1.45 (0.37)	(1.16, 1.80)
	$40,000-$59,999	1.23 (0.21)	(1.02, 1.49)
	$60,000-$79,999	1.04 (0.04)	(0.84, 1.30)
	$80,000+	Ref	
	Missing	0.90 (-0.10)	(0.77, 1.06)
Employment status	Unemployed	2.32 (0.84)	(1.99, 2.70)
Immigrant status	Immigrant	0.85 (-0.16)	(0.71, 1.02)
Constant		0.04 (-3.29)	(0.03, 0.05)

[a] Odds ratio; [b] Beta coefficient; [c] Confidence interval; * Referent category

Table 19
Logistic regression of diagnosis of an anxiety disorder on driving clusters and sociodemographic characteristics ($p < .05$)

Independent variables		OR^{a} $(B)^{b}$	CI^{c} (95%)
K-means cluster	Average Drivers	Ref*	
	Cautious Drivers	1.01 (0.01)	(0.84, 1.20)
	Beltless Drivers	1.60 (0.47)	(1.19, 2.16)
	Egocentric Drivers	1.06 (0.05)	(0.82, 1.36)
	Poly-risk Drivers	1.07 (0.07)	(0.88, 1.30)
Age	<=20	Ref	
	21-25	1.73 (0.55)	(1.25, 2.39)
	26-40	1.59 (0.46)	(1.22, 2.07)
	41-55	1.82 (0.60)	(1.38, 2.38)
	56-70	1.30 (0.27)	(0.94, 1.82)
	71+	0.54 (-0.61)	(0.39, 0.76)
Sex	Male	0.53 (-0.64)	(0.46, 0.61)
Location	Urban	1.06 (0.06)	(0.91, 1.23)
Race	White	Ref	
	Non-White	0.51 (-0.68)	(0.33, 0.78)
	Missing	1.22 (0.20)	(0.94, 1.59)
Marital status	Unmarried	0.62 (-0.47)	(0.53, 0.74)
Education	Less than secondary	0.96 (-0.04)	(0.74, 1.25)
	Secondary graduate	0.97 (-0.03)	(0.78, 1.20)
	Some post-secondary	1.15 (0.14)	(0.87, 1.51)
	Post-secondary graduate	Ref	
	Missing	0.99 (-0.01)	(0.68, 1.46)
Income	$0-$19,999	2.51 (0.92)	(1.88, 3.36)
	$20,000-$39,999	1.42 (0.35)	(1.12, 1.81)
	$40,000-$59,999	1.40 (0.34)	(1.10, 1.78)
	$60,000-$79,999	1.13 (0.13)	(0.88, 1.46)
	$80,000+	Ref	
	Missing	1.06 (0.05)	(0.85, 1.31)
Employment status	Unemployed	1.80 (0.59)	(1.51, 2.16)
Immigrant status	Immigrant	0.63 (-0.46)	(0.49, 0.81)
Constant		0.04 (-3.13)	(0.03, 0.06)

[a] Odds ratio; [b] Beta coefficient; [c] Confidence interval; * Referent category

Table 20
Logistic regression of high levels of stress on driving clusters and sociodemographic characteristics ($p < .05$)

Independent variables		OR^a $(B)^b$	CI^c (95%)
K-means cluster	Average Drivers	Ref*	
	Cautious Drivers	0.84 (-0.18)	(0.77, 0.91)
	Beltless Drivers	1.32 (0.28)	(1.10, 1.58)
	Egocentric Drivers	1.36 (0.30)	(1.20, 1.53)
	Poly-risk Drivers	1.40 (0.34)	(1.26, 1.56)
Age	<=20	Ref	
	21-25	1.34 (0.29)	(1.13, 1.59)
	26-40	1.48 (0.40)	(1.28, 1.72)
	41-55	1.74 (0.56)	(1.50, 2.02)
	56-70	0.93 (-0.07)	(0.80, 1.07)
	71+	0.61 (-0.50)	(0.52, 0.71)
Sex	Male	0.84 (-0.18)	(0.78, 0.90)
Location	Urban	1.08 (0.07)	(1.00, 1.16)
Race	White	Ref	
	Non-White	0.79 (-0.24)	(0.69, 0.90)
	Missing	1.04 (0.04)	(0.90, 1.21)
Marital status	Unmarried	0.98 (-0.02)	(0.91, 1.07)
Education	Less than secondary	1.07 (0.07)	(0.94, 1.22)
	Secondary graduate	0.82 (-0.19)	(0.73, 0.93)
	Some post-secondary	1.28 (0.24)	(1.09, 1.50)
	Post-secondary graduate	Ref	
	Missing	1.07 (0.07)	(0.93, 1.24)
Income	$0-$19,999	1.38 (0.32)	(1.12, 1.70)
	$20,000-$39,999	0.94 (-0.06)	(0.83, 1.08)
	$40,000-$59,999	0.87 (-0.14)	(0.77, 0.98)
	$60,000-$79,999	0.90 (-0.11)	(0.79, 1.01)
	$80,000+	Ref	
	Missing	0.98 (-0.02)	(0.90, 1.08)
Employment status	Unemployed	0.69 (-0.38)	(0.63, 0.75)
Immigrant status	Immigrant	0.99 (-0.01)	(0.90, 1.10)
Constant		1.87 (0.63)	(1.60, 2.18)

[a] Odds ratio; [b] Beta coefficient; [c] Confidence interval; * Referent category

Table 21
Logistic regression of negative (fair and poor) mental health on driving clusters and sociodemographic characteristics ($p < .05$)

Independent variables		OR^{a} $(B)^{b}$	CI^{c} (95%)
K-means cluster	Average Drivers	Ref*	
	Cautious Drivers	0.98 (-0.02)	(0.81, 1.18)
	Beltless Drivers	39.51 (3.68)	(30.73, 50.80)
	Egocentric Drivers	1.06 (0.06)	(0.83, 1.36)
	Poly-risk Drivers	2.42 (0.88)	(1.99, 2.93)
Age	<=20	Ref	
	21-25	1.46 (0.38)	(1.07, 1.98)
	26-40	1.72 (0.54)	(1.32, 2.23)
	41-55	2.52 (0.93)	(1.95, 3.26)
	56-70	1.42 (0.35)	(1.10, 1.83)
	71+	1.24 (0.21)	(0.97, 1.57)
Sex	Male	0.98 (-0.02)	(0.86, 1.11)
Location	Urban	1.01 (0.01)	(0.88, 1.17)
Race	White	Ref	
	Non-White	1.31 (0.27)	(1.05, 1.62)
	Missing	0.26 (-1.35)	(0.16, 0.42)
Marital status	Unmarried	0.75 (-0.29)	(0.65, 0.86)
Education	Less than secondary	1.16 (0.15)	(0.90, 1.48)
	Secondary graduate	1.39 (0.33)	(1.15, 1.69)
	Some post-secondary	1.07 (0.07)	(0.86, 1.34)
	Post-secondary graduate	Ref	
	Missing	0.93 (-0.07)	(0.62, 1.40)
Income	$0-$19,999	2.99 (1.10)	(2.22, 4.03)
	$20,000-$39,999	2.13 (0.76)	(1.69, 2.70)
	$40,000-$59,999	1.23 (0.21)	(0.99, 1.54)
	$60,000-$79,999	1.05 (0.05)	(0.82, 1.35)
	$80,000+	Ref	
	Missing	1.31 (0.27)	(1.10, 1.55)
Employment status	Unemployed	2.75 (1.01)	(2.35, 3.23)
Immigrant status	Immigrant	1.01 (0.01)	(0.86, 1.18)
Constant		0.02 (-4.17)	(0.01, 0.02)

[a] Odds ratio; [b] Beta coefficient; [c] Confidence interval; * Referent category

5.9 References

Anderson, L. R., & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics, 27*(5), 1260-1274.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893.

Bhatti, J. A., Constant, A., Salmi, L. R., Chiron, M., Lafont, S., Zins, M., & Lagarde, E. (2008). Impact of retirement on risky driving behavior and attitudes towards road safety among a large cohort of French drivers (the GAZEL cohort). *Scandinavian Journal of Work, Environment & Health, 34*(4), 307.

Butters, J. E., Mann, R. E., & Smart, R. G. (2006). Assessing road rage victimization and perpetration in the Ontario adult population: The impact of illicit drug use and psychiatric distress. *Canadian Journal of Public Health, 97*(2), 96.

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., & Newstead, S. (2003). Self-regulatory behaviours of older drivers. *Annual proceedings/Association for the Advancement of Automotive Medicine,* Association for the Advancement of Automotive Medicine. *47* 181.

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., Newstead, S., Koppel, S., & O'Hare, M. (2006). Characteristics of older drivers who adopt self-regulatory driving behaviours. *Transportation Research Part F: Traffic Psychology and Behaviour,* 9(5), 363.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association,* 9(3), 522-550.

Dula, C. S., Adams, C. L., Miesner, M. T., & Leonard, R. L. (2010). Examining relationships between anxiety and dangerous driving. *Accident Analysis & Prevention, 4*2(6), 2050-2056.

Evans, L. (2004). In Ferraro C. N. (Ed.), *Traffic safety*. Hauppauge, N.Y.: Nova Science.

Fong, G., Frost, D., & Stansfeld, S. (2001). Road rage: A psychiatric phenomenon*?* *Social Psychiatry and Psychiatric Epidemiology, 3*6(6), 277.

Freeman, J., Maxwell, J. C., & Davey, J. (2011). Unraveling the complexity of driving while intoxicated: A study into the prevalence of psychiatric and substance abuse comorbidity. *Accident Analysis & Prevention, 4*3(1), 34-39.

Ge, Y., Qu, W., Jiang, C., Du, F., Sun, X., & Zhang, K. (2014). The effect of stress and personality on dangerous driving behavior among Chinese drivers. *Accident Analysis & Prevention, 7*3, 34.

Gilmore, H. (2014). Positive mental health and mental illness. Retrieved from http://www.statcan.gc.ca/pub/82-003-x/2014009/article/14086-eng.htm

Government of Canada. (2006). *The human face of mental health and mental illness in Canada*. (No. Cat. No. HP5-19/2006E). Ottawa (Ontario): Minister of Public Works and Government Services Canada. doi:ISBN: 0-662-43887-6

Hindmarch, I., Alford, C., Barwell, F., & Kerr, J. S. (1992). Measuring the side effects of psychotropics: The behavioural toxicity of antidepressants. *Journal of Psychopharmacology,* 6(2), 198.

Hu, T., Xie, X., & Li, J. (2013). Negative or positive? the effect of emotion and mood on risky driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 1*6, 29-40.

Jessor, R. (1987). Risky driving and adolescent problem behaviour: Theoretical and empirical linkage. *T Benjamin (Ed.), Young Drivers Impaired by Alcohol and Other Drugs, Royal Society of Medicine Services, London,* 97-110.

Jonah, B. A. (1986). Accident risk and risk-taking behaviour among young drivers. *Accident Analysis & Prevention, 1*8(4), 255-271.

Kessler, R. C., & Wang, P. S. (2008). The descriptive epidemiology of commonly occurring mental disorders in the United States. *Annual Review of Public Health, 29*, 115-129.

Lajunen, T., & Summala, H. (2003). Can we trust self-reports of driving? effects of impression management on driver behaviour questionnaire responses. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(2), 97-107.

Malta, L. S., Blanchard, E. B., & Freidenberg, B. M. (2005). Psychiatric and behavioral problems in aggressive drivers. *Behaviour Research and Therapy, 43*(11), 1467-1484.

Mann, R. E., Asbridge, M., Stoduto, G., Smart, R. G., Goldbloom, D. S., Vingilis, E. R., & Wickens, C. M. (2010). Psychological distress and collision involvement among adult drivers. *Stress and Health, 26*(2), 127.

Matthews, G. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues in Ergonomics Science, 3*(2), 195.

Mazure, C. M. (1998). Life stressors as risk factors in depression. *Clinical Psychology: Science and Practice, 5*(3), 291.

McDonald, C. C., Sommers, M. S., & Fargo, J. D. (2014a). Risky driving, mental health, and health-compromising behaviours: Risk clustering in late adolescents and adults.*0*(1), 1. doi:10.1136/injuryprev-2014-041150

Murphy, J. M., Horton, N. J., Monson, R. R., Laird, N. M., Sobol, A. M., & Leighton, A. H. (2014). Cigarette smoking in relation to depression: Historical trends from the Stirling County study. *American Journal of Psychiatry, 160*, 1663.

Oltedal, S., & Rundmo, T. (2006). The effects of personality and gender on risky driving behaviour and accident involvement. *Safety Science, 44*(7), 621.

Petridou, E., & Moustaki, M. (2000). Human factors in the causation of road traffic crashes. *European Journal of Epidemiology, 16*(9), 819-826.

Petridou, E., Zavitsanos, X., Dessypris, N., Frangakis, C., Mandyla, M., Doxiadis, S., & Trichopoulos, D. (1997). Adolescents in high-risk trajectory: Clustering of risky behavior and the origins of socioeconomic health differentials. *Preventive Medicine, 26*(2), 215-219.

Rajalin, S. (1994). The connection between risky driving and involvement in fatal accidents. *Accident Analysis & Prevention, 26*(5), 555-562.

Ramaekers, J. G., Ansseau, M., Muntjewerff, N. D., Sweens, J. P., & O'Hanlon, J. F. (1997). Considering the P450 cytochrome system as determining combined effects of antidepressants and benzodiazepines on actual driving performance of depressed outpatients. *International Clinical Psychopharmacology, 12*(3), 159.

Rowden, P., Matthews, G., Watson, B., & Biggs, H. (2011). The relative impact of work-related stress, life stress and driving environment stress on driving outcomes. *Accident Analysis & Prevention, 43*(4), 1332-1340.

Sbarra, D. A., Emery, R. E., Beam, C. R., & Ocker, B. L. (2014). Marital dissolution and major depression in midlife A propensity score analysis. *Clinical Psychological Science, 2*(3), 249.

Scott, D., & Happell, B. (2011). The high prevalence of poor physical health and unhealthy lifestyle behaviours in individuals with severe mental illness. *Mental Health Nursing, 32*(9), 589.

Smart, R. G., Asbridge, M., Mann, R. E., & Adlaf, E. M. (2003). Psychiatric distress among road rage victims and perpetrators. *Canadian Journal of Psychiatry, 48*(10), 681-688.

StataCorp. (2011). In College Station, TX: StataCorp LP. (Ed.), *Stata statistical software: Release 12.*

Statistics Canada. (2011). *Canadian community health survey (CCHS) - annual component [microdata].* Ottawa, Ontario: Data Liberation Initiative [distributor].: Statistics Canada. Retrieved from http://www23.statcan.gc.ca/imdb-bmdi/document/3226_D7_T9_V8-eng.htm

Stephens, A. N., & Groeger, J. A. (2009). Situational specificity of trait influences on drivers' evaluations and driving behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*(1), 29.

Stoduto, G., Dill, P., Mann, R. E., Wells-Parker, E., Toneatto, T., & Shuggi, R. (2008). Examining the link between drinking-driving and depressed mood. *Journal of Studies on Alcohol and Drugs, 69*(5), 777-780.

Taylor, J. E., Deane, F. P., & Podd, J. (2008). The relationship between driving anxiety and driving skill: A review of human factors and anxiety-performance theories to clarify future research needs. *New Zealand Journal of Psychology, 37*(1), 28.

Transport Canada. (2013b). Canadian motor vehicle traffic collision statistics: 2010. Retrieved from http://www.tc.gc.ca/eng/motorvehiclesafety/tp-1317.htm#11

Turner, C., McClure, R., & Pirozzo, S. (2004). Injury and risk-taking behavior—A systematic review. *Accident Analysis & Prevention, 36*(1), 93-101.

Ulleberg, P. (2002). Personality subtypes of young drivers. relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. *Transportation Research Part F: Traffic Psychology and Behaviour,* 4(4), 279-297.

Useche, S., Serge, A., & Alonso, F. (2015). Risky behaviors and stress indicators between novice and experienced drivers. *American Journal of Applied Psychology,* 3(1), 11.

Vaughn, M. G., Define, R. S., DeLisi, M., Perron, B. E., Beaver, K. M., Fu, Q., & Howard, M. O. (2011). Sociodemographic, behavioral, and substance use correlates of reckless driving in the United States: Findings from a national sample. *Journal of Psychiatric Research, 4*5(3), 347.

Vingilis, E., & Wilk, P. (2010). Self-reported motor vehicle injury prevention strategies, risky driving behaviours, and subsequent motor vehicle injuries: Analysis of Canadian national population health survey. *Canadian Public Policy, 3*6, 69-80.

West, R., French, D., Kemp, R., & Elander, J. (1993). Direct observation of driving, self reports of driver behaviour, and accident involvement. *Ergonomics, 3*6(5), 557-567.

Westerman, S. J., & Haigney, D. (2000). Individual differences in driver stress, error and violation. *Personality and Individual Differences, 2*9(5), 981.

Wickens, C. M., Mann, R. E., Stoduto, G., Ialomiteanu, A., Smart, R. G., & Rehm, J. (2013). The impact of probable anxiety and mood disorder on self-reported collisions: A population study. *Journal of Affective Disorders, 14*5(2), 253-255.

Wickens, C. M., Smart, R. G., & Mann, R. E. (2014). The impact of depression on driver performance. *International Journal of Mental Health and Addiction, 1*2(4), 524.

Williams, J., Tregear, S., & Amana, A. (2011). Psychiatric disorders and driver safety: A systematic review. Paper presented at the *Driving Assessment 2011: 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.*

Yu, J., Evans, P. C., & Perfetti, L. (2004). Road aggression among drinking drivers: Alcohol and non-alcohol effects on aggressive driving and road rage. *Journal of Criminal Justice, 3*2(5), 421.

Ziedonis, D., Hitsman, B., Beckham, J. C., Zvolensky, M., Adler, L. E., Audrain-McGovern, J., & ... Riley, W. T. (2008). Tobacco use and cessation in psychiatric disorders: National institute of mental health report. *Nicotine & Tobacco Research, 1*0(12), 1691.

Chapter 6: Thesis Summary And Conclusion

6.0 Thesis Overview

This thesis had two main goals. The first goal was to identify the prevalence and factors associated with six RDBs and to use cluster analysis to assess the homogeneity hypothesis, which examines whether drivers who engage in specific RDBs are distinct from other risky drivers, or whether are they part of a larger profile of drivers who engage in multiple RDBs (Manuscript 1). The second goal was to examine the similarities and differences across mental health variables associated with the subgroups of risky drivers (Manuscript 2).

Secondary data was drawn from the *Driving and Safety* optional module of the CCHS 2009-2011 survey and included 47,356 drivers, ages 16 and older from Newfoundland, Ontario, Alberta, and the Yukon. The prevalence of each RDB in the study population was calculated; the most prevalent RDB was fatigued driving, and the least prevalent RDB was seat belt non-compliance. Such findings were generally consistent with the literature (Traffic Injury Research Foundation, 2014; Transport Canada, 2014; Vanlaar, Simpson, Mayhew, & Robertson, 2008a; Vanlaar, Simpson, Mayhew, & Robertson, 2008b; Vanlaar et al., 2012; CCMTA - STRID sub-group on fatigue).

Twenty-one driver characteristics were distributed across the subcategories of RDBs. Chi-square and t-tests demonstrated associations between driver characteristics and self-reported RDBs, with significant differences ($p < .001$) between all categorical variables and between the means of continuous variables. Associations were apparent between each RDB and male sex, young to middle age, higher SES, White race, poor mental health and engagement in other risk-taking behaviours, and were largely consistent with the literature (Anderson 2008; Asbridge, 2003; Transport Canada, 2013; Wickens, 2011; Wickens, 2014). The results for seat belt non-compliance deviated from this trend; higher proportions of drivers with low SES and who were over the age of 70 reported never wearing a seat belt, compared to drivers with higher SES and who were in younger age categories, respectively. Furthermore, although the literature has indicated associations of young age with both speeding and DUIA, this study demonstrated that these RDBs were also common among middle-aged drivers, indicating that current efforts aimed to reduce such RDBs may also be beneficial to middle-aged drivers (Transport Canada, 2013).

Associations were also found between factors related to health and RDBs, complimenting previous research that suggested links between poor mental health and risky driving (Smart et al., 2003; Wickens et al., 2014; Wickens et al., 2013). Associations were noted between stress and all RDBs, aside from seat belt non-compliance; between diagnosis of an anxiety disorder and all RDBs except DUIA; and between negative mental health as well as diagnosis of a mood disorder with seat belt non-compliance, fatigued driving, speeding, and aggression. In light of previous evidence that different risk taking behaviours tend to coexist in the same individuals, this study included factors representing a propensity to take risks in general (binge drinking, RWDD, and number of injuries in the previous 12 months) (Anderson & Mellor, 2008). This trend was also reflected in the present study. Compared to drivers who refrained from general risk-taking behaviours, drivers who engaged in them were more likely to have reported engagement in all RDBs, aside from seat belt non-compliance.

Acknowledging that many of the individual RDBs shared demographic and behavioural factors common factors, cluster analysis of drivers' engagement in each of six RDBs would reveal whether multiple driving behaviours coexisted among the same drivers, and whether there were different patterns of RDBs.

6.1 The Homogeneity of RDBs

The focus of Manuscript 1 was the cluster analysis of the six RDBs. Cluster analysis revealed five subgroups of drivers in the sample population and included two subgroups of very risky drivers, one subgroup of moderately risky drivers, a small subset of seat belt non-compliant drivers, and one subgroup of cautious drivers. Three patterns of RDBs were evident among the subgroups of risky drivers: 1. The Poly-risk Drivers (20.6%) engaged in the riskiest pattern of driving, including extreme speeding and aggressive driving in addition to driving while fatigued; 2. The Egocentric Drivers (11.7%) engaged in the second riskiest pattern of cell phone-distracted driving and DUIA in addition to moderate levels of speeding and aggression; and 3. The Average Drivers (30.0%) engaged in a moderately risky pattern of occasional cell phone-distracted driving and fatigued driving. The remaining two subgroups included a small subset of drivers, 4. The Beltless Drivers (4.4%), who refrained from all RDBs aside from seat belt non-compliance, and 5. The Cautious Drivers (33.3%) who refrained from all forms of RDBs.

The characteristics associated with the drivers in each cluster further differentiated the five subgroups. Each subgroup (cluster) was uniquely profiled according to a broad range of factors. The profile of the Poly-risk Drivers included male sex, middle age, high income and educational attainment, White race, poor mental health, and engagement in other risk-taking behaviours. The characteristics that profiled the

Egocentric Drivers included young age, male sex, high income and educational attainment, employment, White race, smoking, consumption of alcohol, higher levels of stress, and engagement in other risk-taking behaviours. As per the name assigned to them, the distributions of many of the descriptive characteristics within the Average Drivers cluster were comparable to the sample population – for example, these drivers were generally middle-aged. However, the Average Drivers subgroup featured a higher proportion of post-secondary school graduates than the sample population. The profile of the Beltless Drivers included being older, unmarried, unemployed, of lower educational attainment and income, poorer mental health, and having lower self-perceived health. Finally, the profile of the Cautious Drivers included being female, having a higher mean age, lower educational attainment, lower income, positive mental health, and being more likely than the sample population to be unemployed, non-White, and non-risk takers.

Findings associated with the Beltless Drivers were somewhat different from previous research, which had indicated that seat belt compliance increases with age (Sahai et al., 1998). This subgroup had the highest value for mean age and a large proportion of the cluster members were aged 56 years and older. A number of possible explanations for these findings were considered, but could not be investigated in this study, due to the absence of measures of retirement status, marital dissolution, and year of licensure, as well as the high proportions of missing data for unemployment, education, income, and race, compared to the other four subgroups of drivers.

6.2 Mental Health Factors

Recognizing the associations between the mental health factors and the five clusters found in Manuscript 1, the aim of Manuscript 2 was to explore the four mental health-related variables and their association with cluster membership, examining how poor mental health was associated with specific forms of RDBs. Four separate adjusted logistic regression analyses were performed, regressing each mental health variable on driving clusters and sociodemographic variables ($p < .05$).

Results demonstrated that the risky subgroups of drivers were more likely to have responded positively to the measures of poorer mental health than safer subgroups of drivers. Relative to the Average Drivers, the odds of a diagnosis of a mood disorder, stress, and negative mental health were higher among the Poly-risk Drivers. The odds of reporting higher levels of stress were also higher among the Egocentric Drivers compared to the Average Drivers, while a protective effect was noted for the subgroup of Cautious Drivers. These findings support and extend previous research demonstrating links of psychiatric disorders, higher levels of stress, and negative mental health with dangerous driving behaviours, particularly aggressive driving and speeding (Malta et al., 2005; Rowden et al., 2011; Smart et al., 2003; Wickens et al., 2014; Williams et al., 2011).

Of particular note were the associations between the Beltless Drivers and all four measures of mental health. Specifically, findings of associations with negative mental health suggested that the drivers in this cluster had almost 40 times the odds of reporting negative mental health, compared to drivers in the Average Cluster. These findings highlight a need for further research on this subgroup of drivers.

The results of this study demonstrated two main links between mental health factors and driving behaviour: 1. The diagnosis of a mood disorder, higher levels of stress, and negative mental health were associated with a pattern of severely aggressive driving, speeding and fatigued driving (the Poly-risk Drivers); and 2. Higher levels of stress were associated with drivers who engage in cell phone-distracted driving, DUIA, and moderate levels of aggressive driving and speeding (the Egocentric Drivers). It was noted that further research is required to clarify the contribution of mental health factors to the third pattern of cell phone-distracted driving and fatigued driving (the Average Drivers).

6.3 Limitations and Considerations

Further research is required to confirm the findings of this study, due to the exploratory nature of cluster analysis. Other limitations associated with the present study include the use of unstandardized variables for the clustering procedure. This study did not standardize the variables prior to the clustering procedure due to the small range of variable scales and to avoid masking natural patterns of RDBs in the data (Bible et al., 2013; Milligan & Cooper, 1998). Future research may benefit from comparing the cluster results from both unstandardized and standardized data.

Although the findings of this study draw on driver reports of risky driving, research shows self-reported driving behaviour is a reliable measure of driving behaviour (Lajunen & Summala, 2003; West, French, Kemp, & Elander, 1993). Additionally, although the impairment associated with hands-free cell phone-distracted driving is comparable to hand-held devices, hands-free cell phone use while driving was excluded from this study due to variation in driver attitudes, perceptions of risk, and motivations to use hands-free devices as a safer alternative to hand-held devices (White, Eiser, & Harris, 2004; Zhou, Wu, Rau, &Zhang, 2009).

Finally, although all other health and risk taking-related variables were coded with 0 representing the absence of the characteristic and 1 representing the presence of the characteristic, the variable *satisfaction of life* was measured according to a ten-point scale and coded 1-10, with 1 representing low life satisfaction and 10 representing high life satisfaction. Reverse coding of this variable may have improved clarity and interpretation of the results.

6.4 Implications

Traffic safety research has traditionally focussed on driving behaviour and its impact on crash risk, or has explored a limited number of RDBs, or RDBs independently from one another, often within subpopulations not generalizable to the overall driving population. Although, due to the non-inferential nature of cluster analysis, further research is required to confirm the findings of this study, this study contributes to traffic safety research by confirming the homogeneity hypothesis of RDBs using population data, a broad range of RDBs, and drivers of all ages. The five heterogeneous subsets of drivers revealed by cluster analysis each had a high degree of within-cluster homogeneity and were uniquely profiled by sociodemographic, health, and other risk-taking behaviours. Furthermore, this study highlighted the contribution of mental health factors to dangerous driving. These findings may encourage future research to explore RDBs as homologous risk-taking behaviours, co-existing as patterns of driving behaviours among risky drivers.

While many of the current RDB countermeasures in place must not be discounted, the profiles of the five subgroups of drivers in this study suggest that factors associated with lifestyle, health, culture, and a propensity to engage in risk-taking in general are linked to risky driving. A change in traffic safety culture may be necessary to effectively reduce the prevalence of RDBs. For example, the profile associated with the Egocentric

Drivers included higher SES, young age, unmarried, employed, smoking, urban locations, and higher levels of stress. Efforts to reduce risky driving among these drivers (cell phone-distracted driving, DUIA, and moderate levels of speeding and aggressive driving) may benefit from employment-based interventions to reduce stress, reduce expectations to take work on the road, and to alter the drinking culture of post-work meetings with clients or colleagues. The profile associated with the Poly-risk Drivers consisted of a number of mental health factors such as mood and anxiety diagnoses, higher levels of stress, and negative mental health suggesting that health care workers may play an important role in traffic safety efforts to reduce the number of drivers who engage in a pattern of aggressive driving, speeding, and fatigued driving. As a trusted source of health information, health care professionals may encourage safe driving practices and reinforce traffic safety messages to patients, particularly those with poor mental health. Furthermore, provincial departments of motor vehicles and driver education programs may consider including an educational component which addresses the impact of mental health on driving behaviour and safety, and reinforcing the importance of seat belt compliance when educating, licensing, or re-licensing older drivers.

The literature demonstrates the utility of using methodologies frequently employed in social marketing – such as cluster analysis – as a means to effectively target at-risk populations and promoting change in health-related behaviours such as tobacco use, high-risk sexual behaviours, and physical inactivity (Rovniak, et al., 2010; Spencer, Roberts, Irvine, Jones, & Baker, 2007). As a form of data segmentation, cluster analysis enables researchers to refine their insights about how behavioural changes may most effectively occur, and to more accurately identify potential barriers or contributing factors

(Glanz, Rimer, & Viswanath, 2008). The present study demonstrates that cluster analysis is a viable means to explore driving behaviours and may benefit traffic safety interventions and driver licensing programs, such as the Graduated Driver Licencing (GDL) program, which target their interventions or education to particular populations. For example, interventions or GDL could tailor their driver education based upon driver demographics or program location.

Further research to verify the findings of this study will contribute to a more comprehensive understanding of the role that mental health plays in risky driving, and will enhance traffic safety by enabling better identification of risky drivers.

Bibliography

Aarts, L., & van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention, 38*(2), 215-224.

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis*. Newberry Park, California: Sage University paper series on Quantitative Applications Social Science in the Social Sciences. 07-044.

Anderson, L. R., & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics, 27*(5), 1260-1274.

Asbridge, M., Smart, R. G., & Mann, R. E. (2003). The" homogamy" of road rage: Understanding the relationship between victimization and offending among aggressive and violent motorists. *Violence and Victims, 18*(5), 517-531.

Asbridge, M., Brubacher, J. R., & Chan, H. (2013). Cell phone use and traffic crash risk: A culpability analysis. *International Journal of Epidemiology, 42*(1), 259-267.

Bailey, K. D. (1994). *Typologies and Taxonomies: An Introduction to Classification Techniques.* Sage.

Baird, B. M., Lucas, R. E., & Donnellan, M. B. (2010). Life satisfaction across the lifespan: Findings from two nationally representative panel studies. *Social Indicators Research, 99*(2), 183-203.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893.

Bhatti, J. A., Constant, A., Salmi, L. R., Chiron, M., Lafont, S., Zins, M., & Lagarde, E. (2008). Impact of retirement on risky driving behavior and attitudes towards road safety among a large cohort of French drivers (the GAZEL cohort). *Scandinavian Journal of Work, Environment & Health, 34*(4), 307.

Bible, J., Datta, S., & Datta, S. (2013). Cluster analysis: Finding groups in data. *Informatics for Materials Science and Engineering: Data-Driven Discovery for Accelerated Experimentation and Application,* 53.

Blomberg, R. D., Peck, R. C., Moskowitz, H., Burns, M., & Fiorentino, D. (2009). The long Beach/Fort Lauderdale relative risk study. *Journal of Safety Research, 40*(4), 285-292.

Borkenstein, R. F. (1974). *The role of the drinking driver in traffic accidents: The Grand Rapids study*. Hamburg: Steintor Verlag.

Butters, J. E., Mann, R. E., & Smart, R. G. (2006). Assessing road rage victimization and perpetration in the Ontario adult population: The impact of illicit drug use and psychiatric distress. *Canadian Journal of Public Health, 97*(2), 96.

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*(4), 1282-1293.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-Theory and Methods, 3*(1), 1-27.

Canadian Council of Motor Transport Administrators. (2014). *CCMTA national public opinion survey on road safety: Report on key findings.* (No. 14). Ottawa, Ontario: Canadian Council of Motor Transport Administrators.

Canadian Hospital Injury Report and Prevention Program and Public Health Agency of Canada. (2013). Injury in review, 2012 edition: Spotlight on road and transport safety. Retrieved from http://www.phac-aspc.gc.ca/injury-bles/chirpp/injrep-rapbles/ir-eb-2012-eng.php

CCMTA - STRID sub-group on fatigue. Retrieved from: http://www.ccmta.ca/english/committees/rsrp/strid-fatigue/strid-fatigue-reports.cfm

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., & Newstead, S. (2003). Self-regulatory behaviours of older drivers. *Annual proceedings/Association for the Advancement of Automotive Medicine,* Association for the Advancement of Automotive Medicine. *4*7; 181.

Charlton, J. L., Oxley, J., Fildes, B., Oxley, P., Newstead, S., Koppel, S., & O'Hare, M. (2006). Characteristics of older drivers who adopt self-regulatory driving behaviours. *Transportation Research Part F: Traffic Psychology and Behaviour,* 9(5), 363.

Dahlen, E. R., Martin, R. C., Ragan, K., & Kuhlman, M. M. (2005). Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving. *Accident Analysis & Prevention, 3*7(2), 341-348.

Deery, H. A., & Fildes, B. N. (1999). Young novice driver subtypes: Relationship to high-risk behavior, traffic accident record, and simulator driving performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 41*(4), 628.

Di Milia, L., Smolensky, M. H., Costa, G., Howarth, H. D., Ohayon, M. M., & Philip, P. (2011). Demographic factors, fatigue, and driving accidents: An examination of the published literature. *Accident Analysis & Prevention, 4*3(2), 516-532.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association,* 9(3), 522-550.

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis* Wiley New York.

Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification* John Wiley & Sons.

Dula, C. S., Adams, C. L., Miesner, M. T., & Leonard, R. L. (2010). Examining relationships between anxiety and dangerous driving. *Accident Analysis & Prevention, 4*2(6), 2050-2056.

Elder, R. W., Voas, R., Beirness, D., Shults, R. A., Sleet, D. A., Nichols, J. L., & Compton, R. (2011). Effectiveness of ignition interlocks for preventing alcohol-impaired driving and alcohol-related crashes: A community guide systematic review. *American Journal of Preventive Medicine, 4*0(3), 362-376.

Elvik, R. (2011). Effects of mobile phone use on accident risk. *Transportation Research Record: Journal of the Transportation Research Board, 223*6(1), 20-26.

Evans, L. (1991). *Traffic safety and the drive*r. New York: Van Nostrand Reinhold.

Evans, L. (1996). Safety-belt effectiveness: The influence of crash severity and selective recruitment. *Accident Analysis & Prevention, 2*8(4), 423-433.

Evans, L. (2004). In Ferraro C. N. (Ed.), *Traffic safety*. Hauppauge, N.Y.: Nova Science.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). In Shewhart W. A., Wilks S. S. (Eds.), *Cluster analysis, 5th edition,* Jon Wiley & Sons, Ltd*, Chichester, UK*. (5th ed.).

Fernandes, R., Hatfield, J., & Job, R. (2010). A systematic investigation of the differential predictors for speeding, drink-driving, driving while fatigued, and not wearing a seat belt, among young drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 1*3(3), 179-196.

Fong, G., Frost, D., & Stansfeld, S. (2001). Road rage: A psychiatric phenomenon*? Social Psychiatry and Psychiatric Epidemiology, 3*6(6), 277.

Freeman, J., Maxwell, J. C., & Davey, J. (2011). Unraveling the complexity of driving while intoxicated: A study into the prevalence of psychiatric and substance abuse comorbidity. *Accident Analysis & Prevention, 4*3(1), 34-39.

Gan, G., Ma, C., & Wu, J. (2007). *Data Clustering: Theory, Algorithms, and Applications* (Vol. 20). Siam.

Ge, Y., Qu, W., Jiang, C., Du, F., Sun, X., & Zhang, K. (2014). The effect of stress and personality on dangerous driving behavior among Chinese drivers. *Accident Analysis & Prevention, 7*3, 34.

Gielen, A. C., & Sleet, D. (2003). Application of behavior-change theories and methods to injury prevention. *Epidemiologic Reviews, 2*5, 65-76.

Gilmore, H. (2014). Positive mental health and mental illness. Retrieved from: http://www.statcan.gc.ca/pub/82-003-x/2014009/article/14086-eng.htm

Golias, I., & Karlaftis, M. G. (2001). An international comparative study of self-reported driver behavior. *Transportation Research Part F: Traffic Psychology and Behaviour,* 4(4), 243-256.

Government of Canada. (2006). *The human face of mental health and mental illness in Canada*. (No. Cat. No. HP5-19/2006E). Ottawa (Ontario): Minister of Public Works and Government Services Canada.

Green, L. W., & Kreuter, M. W. (1999). *Health promotion planning: An educational and environmental approach* (3rd ed.). Mountain View, CA: Mayfield Press.

Haddon, W., Jr. (1968). The changing approach to the epidemiology, prevention, and amelioration of trauma: The transition to approaches etiologically rather than descriptively based. *American Journal of Public Health and the Nation's Health, 5*8(8), 1431-1438.

Hastie, T., Tibshirani, R., Friedman, J., (2009). *The Elements of Statistical Learning – Data Mining, Inference, and Prediction.* Springer Series in Statistics.

Hennessy, D. A., & Wiesenthal, D. L. (1997). The relationship between traffic congestion, driver stress and direct versus indirect coping behaviours. *Ergonomics, 4*0(3), 348.

Hindmarch, I., Alford, C., Barwell, F., & Kerr, J. S. (1992). Measuring the side effects of psychotropics: The behavioural toxicity of antidepressants. *Journal of Psychopharmacology,* 6(2), 198.

Hu, T., Xie, X., & Li, J. (2013). Negative or positive? the effect of emotion and mood on risky driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 1*6, 29-40.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health & Social Behavior, 3*8(1), 21-37.

Ivers, R., Senserrick, T., Boufous, S., Stevenson, M., Chen, H. Y., Woodward, M., & Norton, R. (2009). Novice drivers' risky driving behavior, risk perception, and crash risk: Findings from the DRIVE study. *American Journal of Public Health, 9*9(9), 1638-1644.

Iversen, H. (2004). Risk-taking attitudes and risky driving behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour,* 7(3), 135-150.

Jessor, R. (1987). Risky driving and adolescent problem behaviour: Theoretical and empirical linkage. *T Benjamin (Ed.), Young Drivers Impaired by Alcohol and Other Drugs, Royal Society of Medicine Services, London,* 97-110.

Jessor, R., Turbin, M. S., & Costa, F. M. (1997). Predicting developmental change in risky driving: The transition to young adulthood. *Applied Developmental Science,* 1(1), 4.

Jessor, R. (1987). Risky driving and adolescent problem behavior: An extension of problem-behavior theory. *Alcohol, Drugs & Driving,* 3(3-4), 1-11.

Jonah, B. A., & Dawson, N. E. (1987). Youth and risk: Age differences in risky driving, risk perception, and risk utility. *Alcohol, Drugs & Driving.*

Jonah, B. A. (1986). Accident risk and risk-taking behaviour among young drivers. *Accident Analysis & Prevention, 1*8(4), 255-271.

Jonah, B. A. (1990). Age differences in risky driving. *Health Education Research,* 5(2), 139-149.

Juster, F. T., & Smith, J. P. (1997). Improving the quality of economic data: Lessons from the HRS and AHEAD. *Journal of the American Statistical Association, 9*2(440), 1268-1278.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 6*2(6), 593-602.

Kessler, R. C., Olfson, M., & Berglund, P. A. (1998). Patterns and predictors of treatment contact after first onset of psychiatric disorders. *American Journal of Psychiatry, 155*(1), 62-69.

Kessler, R. C., & Wang, P. S. (2008). The descriptive epidemiology of commonly occurring mental disorders in the United States. *Annual Review of Public Health, 29*, 115-129.

Kisely, S., Lin, E., Lesage, A., Gilbert, C., Smith, M., Campbell, L., & Vasiliadis, H. (2009). Use of administrative data for the surveillance of mental disorders in 5 provinces. *Canadian Journal of Psychiatry*, *54*(8), 571.

Lajunen, T., & Parker, D. (2001). Are aggressive people aggressive drivers? A study of the relationship between self-reported general aggressiveness, driver anger and aggressive driving. *Accident Analysis & Prevention, 33*(2), 243-255.

Lajunen, T., & Summala, H. (2003). Can we trust self-reports of driving? effects of impression management on driver behaviour questionnaire responses. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(2), 97-107.

Linden, R., Mann, R. E., Smart, R. G., Vingilis, E., Solomon, R., Chamberlain, E., . . . Stoduto, G. (2010). Research, policy development, and progress: Antisocial behaviour and the automobile. *Canadian Public Policy, 36*, 81-93.

Lucidi, F., Giannini, A. M., Sgalla, R., Mallia, L., Devoto, A., & Reichmann, S. (2010). Young novice driver subtypes: Relationship to driving violations, errors and lapses. *Accident Analysis & Prevention, 42*(6), 1689-1696.

Lustman, M., Wiesenthal, D. L., & Flett, G. L. (2010). Narcissism and aggressive driving: Is an inflated view of the self a road hazard? Journal of Applied Social Psychology, 40, 1423-1449.

Macdonald, S., Anglin-Bodrug, K., Mann, R. E., & Chipman, M. (2005). Driving while impaired (DWI) by alcohol convictions among alcohol, cocaine, and cannabis clients in treatment. *Traffic Injury Prevention, 6*(3), 207-211.

Malta, L. S., Blanchard, E. B., & Freidenberg, B. M. (2005). Psychiatric and behavioral problems in aggressive drivers. *Behaviour Research and Therapy, 43*(11), 1467-1484.

Mann, R. E., Asbridge, M., Stoduto, G., Smart, R. G., Goldbloom, D. S., Vingilis, E. R., & Wickens, C. M. (2010). Psychological distress and collision involvement among adult drivers. *Stress and Health, 26*(2), 127.

Mann, R. E., Smart, R. G., Stoduto, G., Adlaf, E. M., & Ialomiteanu, A. (2004). Alcohol consumption and problems among road rage victims and perpetrators. *Journal of Studies on Alcohol, 65*(2), 161-168.

Mann, R. E., Stoduto, G., Vingilis, E., Asbridge, M., Wickens, C. M., Ialomiteanu, A., . . Smart, R. G. (2010). Alcohol and driving factors in collision risk. *Accident Analysis & Prevention, 42*(6), 1538-1544.

Martiniuk, A. L., Ivers, R. Q., Glozier, N., Patton, G. C., Senserrick, T., Boufous, S., . . . Woodward, M. (2010). Does psychological distress increase the risk for motor vehicle crashes in young people? findings from the DRIVE study. *Journal of Adolescent Health, 47*(5), 488-495.

Matthews, G. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues in Ergonomics Science, 3*(2), 195.

Matthews, G., Dorn, L., Hoyes, T. W., Davies, D. R., Glendon, A. I., & Taylor, R. G. (1998). Driver stress and performance on a driving simulator. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 40*(1), 136-149.

Mazure, C. M. (1998). Life stressors as risk factors in depression. *Clinical Psychology: Science and Practice, 5*(3), 291.

McCartt, A. T., Mayhew, D. R., Braitman, K. A., Ferguson, S. A., & Simpson, H. M. (2009). *Effects of age and experience on young driver crashes: Review of recent literature.* Traffic Injury Prevention, *10*(3), 209-219.

McDonald, C. C., Sommers, M. S., & Fargo, J. D. (2014). Risky driving, mental health, and health-compromising behaviours: Risk clustering in late adolescents and adults. *0*(1), 1.

McGinnis, J. M., Williams-Russo, P., & Knickman, J. R. (2002). The case for more active policy attention to health promotion. *Health Affairs, 21*(2), 78.

McKenna, F.P., Horswill, M.S. (2006). Risk taking from the participant's perspective: The case of driving and accident risk. *Health Psychology, 25*(2), 163-170.

Miller, J. D., Campbell, W. K., Young, D. L., Lakey, C. E., Reidy, D. E., Zeichner, A., & Goodie, A. S. (2009). Examining the relations among narcissism, impulsivity, and self-defeating behaviors. *Journal of personality*, *77*(3), 761-794.

Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification, 5*(2), 181.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50*(2), 159-179.

Moskowitz, H., & Fiorentino, D. (2000). *A review of the literature on the effects of low doses of alcohol on driving-related skills.* (No. HS-809 028). Washington, DC: US Department of Transportation National Highway Traffic Safety Administration.

Murphy, J. M., Horton, N. J., Monson, R. R., Laird, N. M., Sobol, A. M., & Leighton, A. H. (2014). Cigarette smoking in relation to depression: Historical trends from the Stirling County study. *American Journal of Psychiatry, 160*, 1663.

Musselwhite, C. (2006). Attitudes towards vehicle driving behaviour: Categorising and contextualising risk. *Accident Analysis & Prevention, 38*(2), 324-334.

Nilsson, G. (2004). Traffic safety dimensions and the power model to describe the effect of speed on safety. *Bulletin-Lunds Tekniska Högskola, Inst för Teknik och Samhälle, Lunds Universitet*, *221*.

Nurullah, A. S., Thomas, J., & Vakilian, F. (2013). The prevalence of cell phone use while driving in a Canadian province. *Transportation Research Part F: Traffic Psychology and Behaviour,* (19), 52.

Olfson, M., Kessler, R. C., Berglund, P. A., & Lin, E. (1998). Psychiatric disorder onset and first treatment contact in the United States and Ontario. *American Journal of Psychiatry, 155*(10), 1415-1422.

Oltedal, S., & Rundmo, T. (2006). The effects of personality and gender on risky driving behaviour and accident involvement. *Safety Science, 44*(7), 621.

Peck, R. C., Gebers, M. A., Voas, R. B., & Romano, E. (2008). *The relationship between blood alcohol concentration (BAC), age, and crash risk*

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A. A., Jarawan, E., & Mathers, C. D. (2004). World report on road traffic injury prevention.

Petridou, E., & Moustaki, M. (2000). Human factors in the causation of road traffic crashes. *European Journal of Epidemiology, 16*(9), 819-826.

Petridou, E., Zavitsanos, X., Dessypris, N., Frangakis, C., Mandyla, M., Doxiadis, S., & Trichopoulos, D. (1997). Adolescents in high-risk trajectory: Clustering of risky behavior and the origins of socioeconomic health differentials. *Preventive Medicine, 26*(2), 215-219.

Phillips, D. P., & Brewer, K. M. (2011). The relationship between serious injury and blood alcohol concentration (BAC) in fatal motor vehicle accidents: BAC = 0.01% is associated with significantly more dangerous accidents than BAC = 0.00%. *Addiction, 106*(9), 1614-1622.

Quimby, A., Maycock, G., Palmer, C., & Buttress, S. (1999). *The factors that influence a driver's choice of speed: A questionnaire study*. Transport Research Laboratory.

Rajalin, S. (1994). The connection between risky driving and involvement in fatal accidents. *Accident Analysis & Prevention, 26*(5), 555-562.

Ramaekers, J. G., Ansseau, M., Muntjewerff, N. D., Sweens, J. P., & O'Hanlon, J. F. (1997). Considering the P450 cytochrome system as determining combined effects of antidepressants and benzodiazepines on actual driving performance of depressed outpatients. *International Clinical Psychopharmacology, 12*(3), 159.

Rovniak LS, Sallis JF, Saelens BE, Frank LD, Marshall SJ, Norman GJ, Conway TL, Cain KL, Hovell MF, (2010). Adults' physical activity patterns across life domains: Cluster analysis with replication. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association, 29*(5), 496-505.

Rowden, P., Matthews, G., Watson, B., & Biggs, H. (2011). The relative impact of work-related stress, life stress and driving environment stress on driving outcomes. *Accident Analysis & Prevention, 43*(4), 1332-1340.

Runyan, C. W. (2003). Introduction: Back to the future--revisiting Haddon's conceptualization of injury epidemiology and prevention. *Epidemiologic Reviews, 25*, 60-64.

Sahai, V. S., Pitblado, J. R., Bota, G. W., & Rowe, B. H. (1998). Factors associated with seat belt use: An evaluation from the Ontario health survey. *Canadian Journal of Public Health = Revue Canadienne De Sante Publique, 89*(5), 320-324.

Sbarra, D. A., Emery, R. E., Beam, C. R., & Ocker, B. L. (2014). Marital dissolution and major depression in midlife A propensity score analysis. *Clinical Psychological Science, 2*(3), 249.

Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research, 8*(1), 3.

Schmitt, J. A. J., Wingen, M., Riedel, W. J., & Ramaekers, J. G. (2005). *Effects of depression and antidepressant therapy on driving performances* (No. Immortal Vol deliverable 1.5). Maastricht University: European Committee.

Scott, D., & Happell, B. (2011). The high prevalence of poor physical health and unhealthy lifestyle behaviours in individuals with severe mental illness. *Mental Health Nursing, 32*(9), 589.

Scott-Parker, B., Goode, N., & Salmon, P. (2015). The driver, the road, the rules… and the rest? A systems-based approach to young driver road safety. *Accident Analysis & Prevention, 74*, 297-305.

Shope, J. T. (2006). Influences on youthful driving behavior and their potential for guiding interventions to reduce crashes. *Injury Prevention: Journal of the International Society for Child and Adolescent Injury Prevention, 12 Suppl 1*, i9-14.

Simpson, H., & Mayhew, D. (1992). *Reducing the risks for new drivers: A graduated licensing system for British Columbia.* Ottawa, Ontario: Traffic Injury Research Foundation.

Smart, R. G., & Mann, R. E. (2002). Is road rage a serious traffic problem? *Traffic Injury Prevention, 3*(3), 183-189.

Smart, R. G., Asbridge, M., Mann, R. E., & Adlaf, E. M. (2003). Psychiatric distress among road rage victims and perpetrators. *Canadian Journal of Psychiatry, 48*(10), 681-688.

Sokal, R. R., & Sneath, P. H. A. (1963). *Principles of numerical taxonomy.* San Francisco, CA: W. H. Freeman & Co.

Spencer, L., Roberts, G., Irvine, F., Jones, P., & Baker, C. (2007). Using cluster analysis to explore survey data. *Nurse Researcher, 15*(1), 37-54.

StataCorp. (2011). In College Station, TX: StataCorp LP. (Ed.), *Stata Statistical Software: Release 12.*

Statistics Canada. (2011). *Canadian community health survey (CCHS) - annual component [microdata].* Ottawa, Ontario: Data Liberation Initiative [distributor].: Statistics Canada.

Stephens, A. N., & Groeger, J. A. (2009). Situational specificity of trait influences on drivers' evaluations and driving behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*(1), 29.

Stoduto, G., Dill, P., Mann, R. E., Wells-Parker, E., Toneatto, T., & Shuggi, R. (2008). Examining the link between drinking-driving and depressed mood. *Journal of Studies on Alcohol and Drugs, 69*(5), 777-780.

Svenson, O., & Patten, C. J. (2005). Mobile phones and driving: A review of contemporary research. *Cognition, Technology & Work, 7*(3), 182.

Taylor, J. E., Deane, F. P., & Podd, J. (2008). The relationship between driving anxiety and driving skill: A review of human factors and anxiety-performance theories to clarify future research needs. *New Zealand Journal of Psychology, 37*(1), 28.

Traffic Injury Research Foundation. (2014). *The road safety monitor 2014: Drinking and driving in Canada.* (No. 14L). Ottawa Ontario Canada.

Transport Canada. (2005). *Driver attitude to speeding and speed management: A quantitative and qualitative study - final report.* (No. TP 14756 E). EKOS Research Associates Inc.

Transport Canada. (2008). Smashed: A sober look at drinking and driving. Retrieved from http://www.tc.gc.ca/eng/motorvehiclesafety/tp-tp1535-menu-830.htm

Transport Canada. (2013a). Canadian motor vehicle traffic collision statistics 2011 - TrafficCollisionStatisitcs_2011.pdf Retrieved from http://www.tc.gc.ca/media/documents/roadsafety/TrafficCollisionStatisitcs_2011.pdf

Transport Canada. (2013b). Canadian motor vehicle traffic collision statistics: 2010. Retrieved from http://www.tc.gc.ca/eng/motorvehiclesafety/tp-1317.htm#11

Transport Canada. (2014). Road safety in Canada. Retrieved from http://www.tc.gc.ca/eng/motorvehiclesafety/tp-tp15145-1201.htm

Transport Canada. (2015). Canadian motor vehicle traffic collision statistics - 2012.

Transport Research International Documentation (TRID). (2013). Alcohol-crash problem in Canada: 2010. Retrieved from http://ccmta.ca/en/publications/resources-home/item/alcohol-crash-problem-in-Canada-2010

Turner, C., McClure, R., & Pirozzo, S. (2004). Injury and risk-taking behavior—A systematic review. *Accident Analysis & Prevention, 36*(1), 93-101.

Ulleberg, P., & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. *Safety Science, 4*1(5), 427-443.

Ulleberg, P. (2002). Personality subtypes of young drivers. relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. *Transportation Research Part F: Traffic Psychology and Behaviour,* 4(4), 279-297.

Useche, S., Serge, A., & Alonso, F. (2015). Risky behaviors and stress indicators between novice and experienced drivers. *American Journal of Applied Psychology,* 3(1), 11.

Vanlaar, W., Robertson, R. D., & Marcoux, K. (2008). *The road safety monitor 2007: Excessive speeding. Traffic Injury Research Foundation.*

Vanlaar, W., Simpson, H., Mayhew, D. & Robertson, R. (2007). The road safety monitor 2006: Aggressive driving. Retrieved from http://www.trafficinjuryr..._AggressiveDriving_eng.pdf

Vanlaar, W., Simpson, H., Mayhew, D., & Robertson, R. (2008a). Aggressive driving: A survey of attitudes, opinions and behaviors. *Journal of Safety Research, 3*9(4), 375.

Vanlaar, W., Simpson, H., Mayhew, D., & Robertson, R. (2008b). Fatigued and drowsy driving: A survey of attitudes, opinions and behaviors. *Journal of Safety Research, 3*9(3), 303-309.

Vanlaar, W., Robertson, R., Marcoux, K., Mayhew, D., Brown, S., & Boase, P. (2012). Trends in alcohol-impaired driving in Canada. *Accident Analysis & Prevention, 4*8, 297-302.

Vassallo, S., Smart, D., Sanson, A., Harrison, W., Harris, A., Cockfield, S., & McIntyre, A. (2007). Risky driving among young Australian drivers: Trends, precursors and correlates. *Accident Analysis & Prevention, 3*9(3), 444-458.

Vaughn, M. G., Define, R. S., DeLisi, M., Perron, B. E., Beaver, K. M., Fu, Q., & Howard, M. O. (2011). Sociodemographic, behavioral, and substance use correlates of reckless driving in the United States: Findings from a national sample. *Journal of Psychiatric Research, 4*5(3), 347.

Vingilis, E., & Wilk, P. (2010). Self-reported motor vehicle injury prevention strategies, risky driving behaviours, and subsequent motor vehicle injuries: Analysis of Canadian national population health survey. *Canadian Public Policy, 3*6, 69-80.

Vingilis, E. (2007). Predictors of motor vehicle collision injuries among a nationally representative sample of Canadians. *Traffic Injury Prevention,* 8(4), 411.

Wang, P. S., Angermeyer, M., Borges, G., Bruffaerts, R., Tat Chiu, W., DE Girolamo, G., . . . Ustun, T. B. (2007). Delay and failure in treatment seeking after first onset of mental disorders in the world health organization's world mental health survey initiative. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA),* 6(3), 177-185.

West, R., French, D., Kemp, R., & Elander, J. (1993). Direct observation of driving, self reports of driver behaviour, and accident involvement. *Ergonomics, 3*6(5), 557-567.

Westerman, S. J., & Haigney, D. (2000). Individual differences in driver stress, error and violation. *Personality and Individual Differences, 2*9(5), 981.

White, M. P., Eiser, J. R., & Harris, P. R. (2004). Risk perceptions of mobile phone use while driving. *Risk Analysis, 24*, 323.

Wickens, C. M., Smart, R. G., & Mann, R. E. (2014). The impact of depression on driver performance. *International Journal of Mental Health and Addiction, 12*(4), 524.

Wickens, C. M., Mann, R. E., Stoduto, G., Ialomiteanu, A., Smart, R. G., & Rehm, J. (2013). The impact of probable anxiety and mood disorder on self-reported collisions: A population study. *Journal of Affective Disorders, 145*(2), 253-255.

Williams, J., Tregear, S., & Amana, A. (2011). Psychiatric disorders and driver safety: A systematic review. Paper presented at the *Driving Assessment 2011: 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design,*

Willis, C., Lybrand, S., & Bellamy, N. (2004). Alcohol ignition interlock programmes for reducing drink driving recidivism. *Status and Date: Edited (no Change to Conclusions).*

Wilson, R. (1990). The relationship of seat belt non-use to personality, lifestyle and driving record. *Health Education Research,* 5(2), 175-185.

World Health Organization. (2009). *Global Status Report on Road Safety: Time for Action. Geneva: World Health Organization, 2009,*

Yu, J., Evans, P. C., & Perfetti, L. (2004). Road aggression among drinking drivers: Alcohol and non-alcohol effects on aggressive driving and road rage. *Journal of Criminal Justice, 32*(5), 421.

Zador, P. L. (1991). Alcohol-related relative risk of fatal driver injuries in relation to driver age and sex. *Journal of Studies on Alcohol, 52*(4), 302.

Zhao, N., Reimer, B., Mehler, B., D'Ambrosio, L. A., & Coughlin, J. F. (2013). Self-reported and observed risky driving behaviors among frequent and infrequent cell phone users. *Accident Analysis & Prevention, 61*, 71-77.

Ziedonis, D., Hitsman, B., Beckham, J. C., Zvolensky, M., Adler, L. E., Audrain-McGovern, J., & ... Riley, W. T. (2008). Tobacco use and cessation in psychiatric disorders: National institute of mental health report. *Nicotine & Tobacco Research, 10*(12), 1691.

Zhou, R., Wu, C., Rau, P. -. P., & Zhang, W. (2009). Young driving learners' intention to use a handheld or hands-free mobile phone when driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*(3), 208.

Appendix A: Haddon's Matrix

An example of how Haddon's Matrix Facilitates the Exploration of the Potential Factors Associated with Motor Vehicle Collisions.

Phase	Host (Drivers)	Agent/Vectors (Motor Vehicles & Equipment)	Physical Environment	Social Environment
Pre-crash	• Demographic factors • Health factors • Personality characteristics • Perceived environment • Developmental factors (physical, psychosocial, behavioural) • Police enforcement • Passengers • Driving behaviours	• Vehicle power • Maximum speed • Road-worthiness • Lighting • Braking • Handling • Speed management • Seat belt in-car reminder • BAC testers	• Road design & layout • Weather • Speed limits • Pedestrian facilities • Rumble strips • Traffic calming schemes • Safety signage • Roadway lighting	• Traffic safety culture (attitudes & beliefs) • Community norms • Government policies & laws • Employment environment • Education • Availability of public transport • Establishment regulations
Crash	• Driving behaviours (seat belt non-compliance, excess speed) • Presence of alcohol or drugs • Human tolerance factors	• Occupant restraints • Other safety devices • Crash protective design • Seats for larger children	• Crash-protective roadside objects • Location (urban or rural)	• Availability of emergency services & resources
Post-crash	• Self-help or help from bystanders • Emergency response team (detection & response to crash) • First-aid skills	• Ease of access • Fire • Leakage of hazardous materials • Presence of alcohol or drugs	• Ease of access • Leakage of hazardous materials • Rescue facilities • Traffic congestion • Changes to high-risk crash sites	• Pre-hospital care • Emergency room & hospital care room care • Psychosocial services • Social support • Insurance • Traffic rules & policy • Support for research

Appendix B: Haddon's Matrix Applied To Risky Driving Behaviours

Example of how Haddon's Matrix facilitates the exploration of the potential factors associated with RDBs. This matrix considers the pre-event, event, and post-event of the occurrences of six RDBs (seat belt non-compliance, cell phone-distracted driving, fatigued driving, speeding, aggressive driving, and DUIA).

Phase	Host (Drivers)	Agent/Vectors (Motor vehicles & Equipment)	Physical Environment	Social Environment
Pre-event	• Demographic factors (SES, age, sex, employment) • Physical health & mental health (stress, negative mental health, psychiatric disorders) • Personality characteristics • Impulsivity • Perceptions/attitudes towards risk • Time urgency (stress) • Passengers • Other risk-taking behaviours (smoking & RWDD) • Alcohol dependency or problems	• Vehicle type • Vehicle power • Maximum speed • Cruise control • Lighting • Braking • Handling • Speed management • Bluetooth device • Hands-free cellular phone device • Seat belts & child safety belts	• Road design and layout • Weather • Speed limit signage • Rest stop signage • Rest stop exits • Rumble strips • Traffic calming schemes • Visibility • Traffic congestion	• Traffic safety culture (attitudes and beliefs) • Community norms • Cultural norms • Lifestyle • Police enforcement • Government policies & laws • Establishment regulations • Peer pressure • Speed limits
Event	• Other RDBs (seat belt non-compliance) • Presence of alcohol or drugs • Human tolerance factors	• Speed of vehicle • Size of vehicle	• Crash-protective roadside objects • Location (urban or rural) • Road surface conditions	• Speed limits • Peer pressure
Post-event	• Change in perception of risk • Comorbidities • Mental health • Physical health	• Ignition interlock system • Removal of vehicle cellular devices	• Traffic congestion • Implementation of traffic calming schemes	• Change in traffic safety culture • Countermeasures such as • Regionalized trauma care • Psychosocial services • Social support • Insurance • Traffic rules & policy amendments • Support for research

Appendix C: Cluster Linkage Parameter Values For Clustering Linkage Methods

Clustering linkage method	α_i	α_j	β	γ
Single	1/2	1/2	0	-1/2
Complete	1/2	1/2	0	1/2
Average	$\frac{n_i}{n_i + n_j}$	$\frac{n_i}{n_i + n_j}$	0	0
Weighted average	1/2	1/2	0	0
Centroid	$\frac{n_i}{n_i + n_j}$	$\frac{n_i}{n_i + n_j}$	$-\alpha_i, \alpha_j$	0
Median	1/2	1/2	-1/4	0
Ward's	$\frac{n_i + n_k}{n_i + n_j + n_k}$	$\frac{n_j + n_k}{n_i + n_j + n_k}$	$\frac{-n_k}{n_i + n_j + n_k}$	0

Appendix D: Primary Outcome Variables (2011 Canadian Community Health Survey Driving And Safety Optional Module)

Domain	CCHS Variable	Variable Type	Variable Recoded	Freq. (%)[*]	Dk/ Ref[a]	DNR[b]	N/A[c]	*N*	Population
Seat Belt Non-Compliance	How often do you fasten your seat belt when you drive a motor vehicle?	Ordinal	Always	41,589 (87.8)	12	2,478	77,857	124,870	28,725,105
			Most of the time	2,225 (4.7)					
			Rarely	457 (1.0)					
			Never	252 (0.5)					
Distracted Driving	Excluding hands-free use, how often do you use a cell phone while you are driving a motor vehicle?	Ordinal	Often	2,486 (5.3)	17	2,480	77,857	124,870	28,725,105
			Sometimes	9,518 (20.1)					
			Rarely	16,214 (34.2)					
			Never	15,732 (33.2)					
Fatigue	How often do you drive when you are feeling tired?	Ordinal	Often	3,114 (6.6)	80	2,485	77,732	124,870	28,725,105
			Sometimes	9,518 (20.1)					
			Rarely	16,214 (34.2)					
			Never	15,732 (33.2)					

Domain	CCHS Variable	Variable Type	Variable Recoded	Freq. (%)*	Dk/ Ref[a]	DNR[b]	N/A[c]	*N*	Population
Speeding	Compared to other drivers, would you say you usually drive:	Ordinal	Much faster	616 (1.3)	178	2,489	77,732	124,870	28,725,105
			A little faster	8,733 (18.4)					
			About the same	27,250 (57.5)					
			A little slower	7,383 (15.6)					
			Much slower	489 (1.0)					
Aggressive Driving	Compared to other drivers, would you say you usually drive:	Ordinal	Much more aggressively	377 (0.8)	322	2,494	77,732	124,870	28,725,105
			A little more aggressively	4,915 (10.4)					
			About the same	19,331 (40.8)					
			A little less aggressively	13,256 (28.0)					
			Much less aggressively	6,443 (13.6)					
Impaired Driving	In the past 12 months, have you driven a motor vehicle after having 2 or more drinks in the hour before you drove?	Dichoto-mous	Yes	2,228 (4.7)	39	1,924	86,583	124,870	28,725,105
			No	34,096 (72.0)					

[*] Calculations of frequency percent (%) distributions exclude N/A, [a] Do not know/Refuse - Includes missing data from Quebec (Health Region 2410): n=68, [b] Did not report, [c] Not Applicable

Appendix E: Independent Variables Of Interest (2011 Canadian Community Health Survey)

CCHS Variable	Variable type	Variable categories	Code	Freq (%)*	Dk/Ref[a]	DNR[b]	N/A[c]	*n*	*N*	Population
Demographic variables										
Age	Continuous	<=20	1	16,592 (13.3)	-	-	77,514	47,356	124,870	28,725,105
		21–25	2	9,868 (7.9)						
		26–40	3	29,505 (23.6)						
		41-55	4	33,631 (26.9)						
		56-70	5	23,455 (18.8)						
		71+	6	11,819 (9.5)						
Sex	Dichotomous	Female	0	56,466 (45.2)	-	682 (0.5)	77,514	47,356	124,870	28,725,105
		Male	1	67,722 (54.2)						
Geography	Dichotomous	Rural	0	22,380 (17.9)	-	-	77,514	47,356	124,870	28,725,105
		Urban	1	102,490 (82.1)						
Immigrant Status	Dichotomous	No	0	103,827 (83.2)	-	3,398 (2.7)	77,514	47,356	124,870	28,725,105
		Yes	1	17,645 (14.1)						
Race	Categorical	White	1	104,817 (83.9)	-	3,617 (2.9)	77,514	47,356	124,870	28,725,105
		Non-White	2	16,436 (13.2)						
		Missing	3	3,617 (2.9)						
Marital Status	Categorical	Married or common-law	0	62,916 (50.7)	-	942 (0.8)	77,514	47,356	124,870	28,725,105
		Not married	1	61,012 (49.1)						
Education	Ordinal	Less than secondary	1	31,493 (25.4)	-	4,067(3.3)	77,514	47,356	124,870	28,725,105
		Secondary graduate	2	19,192 (15.5)						
		Some post-secondary	3	8,871 (7.1)						

CCHS Variable	Variable type	Variable categories	Code	Freq (%)*	Dk/Ref[a]	DNR[b]	N/A[c]	*n*	*N*	Population
		Post-secondary graduate	4	61,247 (49.3)						
Income	Ordinal	$0-$19,999	1	12,413 (1.0)	-	22,068 (17.7)	77,514	47,356	124,870	28,725,105
		$20,000-$39,999	2	22,248 (17.9)						
		$40,000-$59,999	3	19,136 (15.4)						
		$60,000-$79,999	4	15,526 (12.5)						
		$80,000 or more	5	33,479 (27.1)						
		Missing	6	22,068 (17.7)						
Employment Status	Dichotomous	Employed	0	73,754 (59.4)	29 (0.0)	2,058 (1.7)	19,664	47,356	124,870	28,725,105
		Unemployed	1	29,365 (23.6)						
Self-perceived Health	Ordinal	Healthy	0	69,624 (56.0)	-	82 (0.7)	77,514	47,356	124,870	28,725,105
		Unhealthy	1	54,420 (43.8)						
Satisfaction With Life	Ordinal	0 very dissatisfied	0	539 (0.4)	-	4,315 (3.5)	77,514	47,356	124,870	28,725,105
		1	1	178 (0.1)						
		2	2	450 (0.4)						
		3	3	776 (0.6)						
		4	4	1,214 (1.0)						
		5	5	6,453 (5.2)						
		6	6	6,123 (4.9)						
		7	7	19,344 (14.5)						
		8	8	39,942 (32.0)						
		9	9	23,490 (18.8)						
		10 Very satisfied	10	26,361 (21.1)						

Mental Health Status	Ordinal	Positive	0	114,186 (94.0)	-	3,362 (2.7)	77,514	47,356	124,870	28,725,105
		Negative	1	7,322 (6.0)						
Stress	Ordinal	Lower levels of stress	0	49,566 (39.7)	479 (0.4)	682 (0.6)	77,514	47,356	124,870	28,725,105
		Higher levels of stress	1	74,143 (59.4)						
Mood Disorder	Dichotomous	No mood	0	114,731 (91.9)	143 (0.1)	702 (0.6)	77,514	47,356	124,870	28,725,105
		Mood	1	9,294 (7.4)						
Anxiety Disorder	Dichotomous	No anxiety	0	116,883 (93.6)	166 (0.1)	702 (0.7)	77,514	47,356	124,870	28,725,105
		Anxiety	1	7,119 (5.7)						
Smoking Status	Ordinal	Non-smoker	0	97,754 (78.7)	44 (0.0)	980 (0.8)	77,514	47,356	124,870	28,725,105
		Smoker	1	26,092 (21.0)						
Alcohol	Dichotomous	No	0	30,326 (24.4)	139 (0.1)	1,677 (1.3)	77,514	47,356	124,870	28,725,105
		Yes	1	92,728 (74.7)						
Binge Drinking	Ordinal	Non-drinker	1	42618 (34.8)	585 (0.5)	1,834 (1.5)	77,514	47,356	124,870	28,725,105
		Drinker, non-binger	2	49,507 (40.4)						
		Binger	3	30,326 (24.8)						
RWDD	Dichotomous	No	0	50,795 (87.8)	403 (0.3)	2,534 (2.0)	66,307	47,356	124,870	28,725,105
		Yes	1	4,831 (8.3)						
Number of Injuries	Ordinal	Not injured	1	106,366(85.2)	182 (0.2)	-	77,514	47,356	124,870	28,725,105
		1 time	2	13,643 (10.9)						
		2 or more times	3	4,861 (3.9)						

* Calculations of frequency percent (%) distributions exclude N/A, [a] Do not know/Refuse to answer, [b] Did not report, [c] Not applicable

Appendix F: Descriptive Statistics Of The Sample Population Independent Variables [Number (Percent), Mean, And Standard Deviation]

Independent Variables		N (%)	M[a]	SD[b]	Total
Sociodemographic variables					
Age	<=20	3,522 (7.4)	3.7	1.3	47,356 (100)
	21–25	4,003 (8.5)			
	26–40	12,423 (26.2)			
	41-55	14,432 (30.5)			
	56-70	9,181 (19.4)			
	71+	3,795 (8.0)			
Sex	Female	22,510 (47.5)	0.5	0.5	47,356 (100)
	Male	24,846 (52.5)			
Geographic Area	Rural	7,859 (16.6)	0.8	0.4	47,356 (100)
	Urban	39,497 (83.4)			
Immigrant Status	No	35,298 (74.5)	0.3	0.4	47,356 (100)
	Yes	12,058 (25.5)			
Race	White	36,863 (77.8)	1.2	0.4	47,356 (100)
	Non-White	7,826 (16.5)			
	Missing	2,667 (5.6)			
Marital Status	Married or common-law	31,137 (65.8)	0.3	0.5	47,356 (100)
	Not married	16,220 (34.3)			
Education	Less than secondary	2,063 (4.4)	3.6	0.9	47,356 (100)
	Secondary graduate	4,513 (9.5)			
	Some post-secondary	2,239 (4.7)			
	Post-secondary graduate	34,809 (73.5)			
	Missing	3,732 (7.9)			
Income	$0-$19,000	1,372 (2.9)	4.00	1.3	47,356 (100)

Independent Variables		*N* (%)	*M*[a]	*SD*[b]	Total
	$20,000-$39,000	3,747 (7.9)			
	$40,000-$59,000	4,691 (9.9)			
	$60,000-$79,000	4,948 (10.5)			
	$80,000 or more	15,887 (33.6)			
	Missing	16,712 (35.3)			
Employment Status	Employed	36,140 (76.3)	0.2	0.4	47,356 (100)
	Unemployed	11,216 (23.7)			
Continuous Variable			*M*	*SE*	
Age2		47,356 (100)	44.8	0.1	47,356 (100)
Health-related variables					
Self-perceived Health	Healthy	42,376 (89.5)	0.1	0.3	47,356 (100)
	Not healthy	4,980 (10.5)			
Continuous Variable			*M*	*SE*	
Satisfaction With Life	0 very dissatisfied	209 (0.4)	8.1	0.0	47,356 (100)
	1	51 (0.1)			
	2	188 (0.4)			
	3	286 (0.6)			
	4	446 (0.9)			
	5	2,360 (5.0)			
	6	2,365 (5.0)			
	7	7,518 (15.9)			
	8	15,336 (32.4)			
	9	8,668 (18.3)			
	10 Very satisfied	9,928 (21.0)			
Mental Health Status	Positive	44,102 (93.1)	0.1	0.3	47,356 (100)
	Negative	3,254 (6.9)			

Independent Variables		N (%)	M^a	SD^b	Total
Stress	Lower levels of stress	15,433 (32.6)	0.7	0.5	47,356 (100)
	Higher levels of stress	31,923 (67.4)			
Mood Disorder	No mood	44,262 (93.5)	0.1	0.3	47,356 (100)
	Mood	3,094 (6.5)			
Anxiety Disorder	No anxiety	45,185 (95.4)	0.1	0.2	47,356 (100)
	Anxiety	2,172 (4.6)			
Smoking Status	Non-smoker	37,674 (79.6)	0.2	0.4	47,356 (100)
	Smoker	9,682 (20.4)			
Alcohol	No	9,456 (20.0)	0.8	0.4	47,356 (100)
	Yes	37,900 (80.0)			
Other risk-taking behaviours					
Binge Drinking	Non-drinker	9,640 (20.4)	2.2	0.8	47,356 (100)
	Drinker, non-binger	18,993 (40.1)			
	Binger	18,723 (39.5)			
RWDD	No	42,589 (89.9)	0.1	0.3	47,356 (100)
	Yes	4,767 (10.1)			
Number of Injuries	Not injured	40,437 (89.9)	1.2	0.5	47,356 (100)
	1 time	5,150 (10.9)			
	2 or more times	1,769 (3.7)			

*Variable mean, standard deviation, and total do not include missing category; [a] Mean; [b] Standard deviation

Appendix G: Duda-Hart And Calinski/Harabasz Stopping Rules Results

Number of Clusters	Duda-Hart stopping rule		Calinski/Harabasz stopping rule
	Je(2)/Je(1)	Psuedo T-squared	Pseudo-F
1	0.2	40962.4	-
2	0.5	4907.4	40,962
3	0.3	5761.8	49,228
4	0.6	1002.3	48,516
5	0.6	630.3	53,610
6	0.2	2591.8	55,789
7	0.4	2325.2	54,928
8	0.7	311.6	55,645
9	0.6	1135.9	58,006
10	0.4	321.5	60,190
11	0.6	229.6	60,470
12	0.3	1024.9	61,264
13	0.4	306.1	63,290
14	0.6	314.4	64,305
15	0.5	962.5	64,640