STRUCTURAL EQUATION MODELING IN ROAD ACCIDENT RESEARCH: A LITERATURE REVIEW

Supamit Srisawat, Pairoj Roathanachonkun
Faculty of Logistics, Burapha University
169 Long-hard Bangsaen Rd. Mueang District, Chonburi, 20131, Thailand
Email: supamits@yahoo.com, pairoj@buu.ac.th

ABSTRACT
Purpose: This study is presented the applications of Structural Equation Modeling (SEM) in Road accident research and identified some significant research gaps for future research.

Design/methodology/approach: This paper is focused on the applications of Structural Equation Modeling (SEM) in Road accident research. A systematic literature review methodology was employed to analyze published applications of the Structural Equation Modeling (SEM) and identify methodological issues gleaned from reviewing those literatures.

Findings: The findings of this paper focus on empirical applications of SEM which include: (1) CFA models; (2) PA models; and (3) SEM models that combine both measurement and structural components. Research studies in the past focused on the association between model and empirical data. However, they lacked of advance statistical analysis of the variables relationship in the model such as Mediator analysis and Moderator analysis etc.

Research limitations/implication: This study was resulted from reviewing literatures related to applications of SEM from the previous publications in the international journals in the past decade. However, this study was excluded other forms of published applications of SEM.

Practical implications: The findings lead to development of SEM model to apply in road accident researches which the developed model is able to describe causes of road accident with more understanding.

Originality/value: The results of this study are presented understanding situation of application of Structural Equation Modeling (SEM) and indicating guideline for road accident research in the future with advance level.

Keywords: Structural Equation Modeling(SEM), Confirmatory Factor Analysis(CFA), Path Analysis(PA), Road accident model, review literature**.**

Introduction

The Structural Equation Models (SEM) is defined as the method for measuring relationships among latent variables which is originated by Sewall Wright's in 1916 (Bollen, 1989). This model was not used pervasively until 1980 when it was brought by Bagozzi(1980). This technique has been attracted many marketers and consumer behavior researchers since then. For the road accidents researchers, this new statistical approach was not popular at the beginning. But recently, structural equation modeling (SEM) has become well known among those empirical researchers who are conducting researches on the road accidents. Furthermore, Structural Equation Models are employed in many related research articles as a tool to analyze the primary data which appear in the various international journals. Although the SEM application of road accident is used regularly and frequently, there are few guidelines and even fewer standards of the SEM application that researchers remain to conduct analyses, presenting and interpreting the results. Hence, there are still large variances in the results across related publications.

For example, the SEM reviews which appear occasionally found in psychology (Hershberger, 2003), marketing (Baumgartner and Homburg, 1996), MIS (Chin and Todd, 1995; Gefen et al., 2000), strategic management (Shook et al., 2004), logistics (Garver and Mentzer, 1999), and organizational research (Medsker et al., 1994). Even though there are various fields of SEM application, there is no review on road accidents. What is more, there still some issues that researcher should realize about it. Steiger (2001) states that there are many SEM textbooks disregard many significant issues. He notices that it could be insufficient guidance about using SEM correctly. Because of the complexities of using SEM and revealed problems in other fields, a specific review of the literature of road accidents seems to take time and has to be warranted. Accordingly, to conduct the research, there are three objectives. Firstly, we differentiate the published research about road accidents in terms of related criteria such as (1) sample size; (2) variables; and (3) model measurements. At the beginning, we define a different when using two conditions which are frequently used interchangeably in road accident covariance structure modeling (CSM) and structural equation modeling (SEM). Long(1983) mentions that CSM represents a common class of models including ARMA time series models, multiplicative models for multi-faceted data, circumplex models, as well as all SEM models, as well as all SEM models.

Accordingly, it could be seen that SEM models are a subset of CSM models. The current review of SEM models is limited by the researcher because other types of CSM models are not often used in road accident research. In addition, Structural Equation Modeling is defined as a tool to specify, estimate, and evaluate models of linear relationships among a set of variables which are observed in terms of using a generally smaller number of unobserved variables. SEM models consist of observed variables which can be called manifest or measured, measurement variables (MVs) for short and unobserved variables (underlying or latent, LVs for short) that can be a nature independent (exogenous) or dependent (endogenous). LVs, hypothetical constructs, cannot be directly measured because multiple MVs naturally represent the SEM which is served to be indicators of the underlying constructs. Moreover, the SEM model is defined as a priori hypothesis about a pattern of linear relationships among a set of observed and unobserved variables.

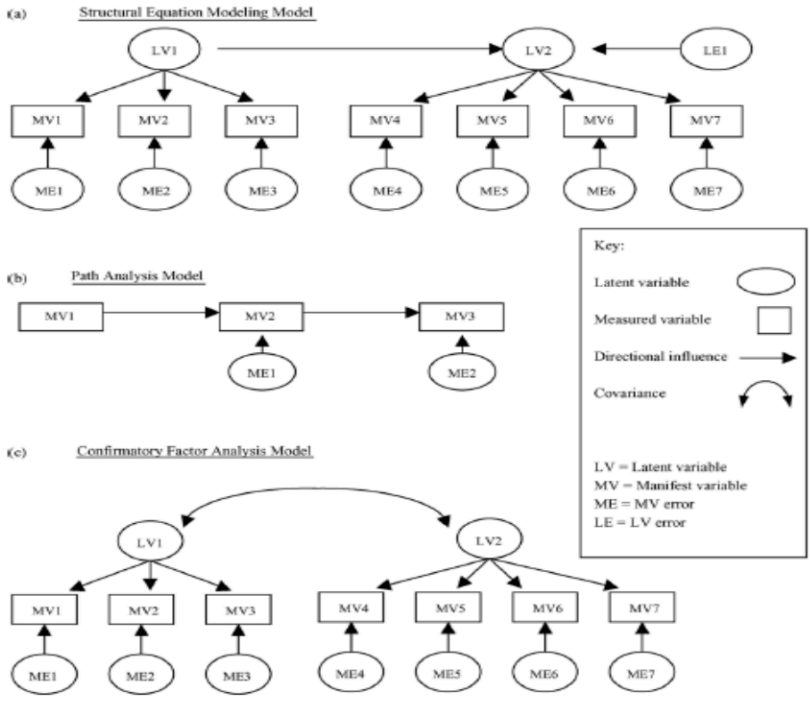

Figure 1 The Presentations of SEM, PA and CFA models (Goldstein and Shah, 2005)

Gefen et al (2000) identify that the main purpose of using SEM is to establish the validity of a priori model rather than to 'find' a suitable model. Therefore, two special cases of SEM that are regularly used in road accident researches are Path Analysis (PA) and Confirmatory Factor Analysis (CFA). The patterns of directional and non-directional relationships are specified by PA models among MVs. Long (1983) notes that to use SEM, researcher has to refer to three model types compost of SEM, PA and CFA as can be seen in figure 1.

Review of published SEM research

In this study, the researcher emphasizes on the articles which have been published internationally under the topic how to apply SEM in the research on the road accidents. The scopes of this research are to study 1) Confirmatory Factory Analysis (CFA models) 2) Path Analysis (PA models) and 3) Structural Equation Models (SEM), which covers both 1) and 2). However, in this research, regression analysis is not included such as Partial Least Squares (PLS) and Exploratory Factor Analysis (EFA models). That is because the analysis of EFA only focuses on measuring variables (MVs) and each group of Latent Variables (LVs) while the scope of SEM is broader than EFA. The main purpose of regression and PLS models are to predict the variance of explanation in the dependent variable(s) which are compared to the development of theory and testing in the form of structural relationships (i.e. parameter estimation) in SEM. The theoretical difference between these approaches is the critical decision whether to use PLS or SEM (Anderson and Gerbing,1988) Additionally, the fundamental assumptions in PLS and regression are less constraining than SEM, therefore, this leads to the problems and concerns in conducting these analyses resulting a significantly different. As a result, regression and PLS models are not included in our review.

Methodology

In this study, a systematic review of literature is used to gain more accurate results and validity. Cook et al. (1997) suggests that to apply a systematic review in the research, it is necessary to explore the systematic way and determine the criteria to select the reports of the previous researches clearly to avoid investigator bias. Moreover, the researcher dedicated to present the results based on the application of SEM in the research on road accidents by using the search term "road accidents" and "Structural Equation Models" in different four main databases including Emerald, SprinkerLink, ProQuest and Science Direct.

Critical issues in the application of SEM

Having said about SEM, there are many significant issues which have been considered whether for evaluating a measurement model or examining the fit of structural relationships, separately or simultaneously. In this research, we concentrate on three issues: (1) sample size; (2) variables; and (3) model measurements including issues related to evaluation, interpretation and presentation of results.

Sample size

Sample size issues have been attracted by several authors (e.g., Anderson and Gerbing, 1988; Bentler, 1990; Bentler and Yuan, 1999; Bollen, 1990; Hoogland and Boomsma, 1998). The mostly commonly used SEM estimation methods today are: normal-theory maximum likelihood (ML), generalized least squares (GLS), weighted least squares (WLS), and etc. However, the ML method is frequently used, which also requires a sufficient sample size, particularly when non-normal data are involved. According to Monte Carlo studies of the performance of various estimation methods, there are several proposed heuristics: (1) A minimum sample size of 200 is necessary to reduce biases to an acceptable level for any type of SEM estimation (Kline, 1998; Loehlin, 1998; Boomsma and Hoogland, 2001). (2) Sample size of ML estimation should be at least fifteen times for number of MVs (Stevens, 1996). (3) Sample size for ML estimation should be at least five times of the number of free parameters in the model, together with error terms (Bentler and Chou, 1987; Bentler, 1995); and (4) with strongly kurtotic data, the minimum sample size should be ten times the number of free parameters (Hoogland and Boomsma, 1998).

Adequacy of sample size has a significant impact on the reliability of parameter estimates, model fit, and statistical power. MacCallum et al (1996) question that how large a sample should be for SEM is deceptively difficult to determine because it is dependent upon several characteristics such as number of MVs per LV, degree of multivariate normality (West et al., 1995), and estimation method (Tanaka, 1987). Suggested approaches for determining sample size include establishing a minimum (e.g., 200), having a certain number of observations per MV, having a certain number of observations per parameters estimated (Bentler and Chou, 1987; Bollen, 1989; Marsh et al., 1988), and through conducting power analysis (MacCallum et al., 1996). While the first two approaches are totally different, the latter two have been studied extensively. According to the previous researches on the road accidents, it found that there are several related articles which use both approaches. For the minimum of sample size, 100 sample sizes is the least number which has been used in the SEM application research. Having discussed about sample size, a certain number of observations per parameters estimated has been use widely in international publication.

Variables in SEM model

SEM models consist of observed variables (Measured variable: MV) and unobserved variables (Latent variable: LV) that can be independent (exogenous) or dependent (endogenous) in nature. LVs are hypothetical constructs that cannot be directly measured, and in SEM are typically represented by multiple MVs that serve as indicators of the underlying constructs. The sources of variables have been taken from the previous adapted models. According to the study of SEM application, it could be categorized by three different groups which based on road safety approaches, behavior and psychology approaches, and the integrated approaches.

For the first group, it based on road safety approaches which consist of human factor, vehicles factor and environmental factor (see figure 2).

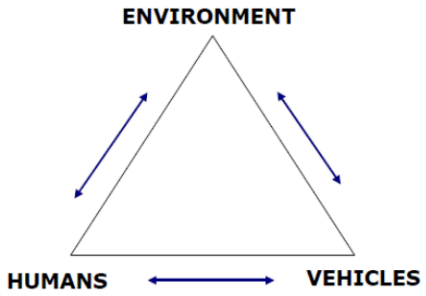

Figure 2 Road Safety Triangle

The triangle concept is based on an intuitive understanding. Among three factors present the relationship in a complex system of road safety. Human-being as those traveling can predefine environment inside or outside moving vehicles. Many different approaches and theories are based on this concept. (e.g. the driver-vehicle interaction(interface), a large area of interest of vehicle manufacturers and psychologists, or vehicle-road interaction, a broad study area of both road and vehicle engineers). Eksler, (2010) agrees that the pyramid concept has been preferred by those researchers and policy makers. The triangle concept has traditionally been used by road safety practitioners and other stakeholders.

For the second group, it based on behavior and psychology approaches, which normally study about the mechanistic of road accidents. In the study of driving behavior, the researcher focuses on the causes of risk-driving, which leads to road accidents such as Risk perception, Attitude, Driving skill, and the mood of Driver. In the study of driving psychology, it has been focused on the decision-making process and the risk perception such as Theory of Planned Behavior (TPB), which is significant principle of these approaches.

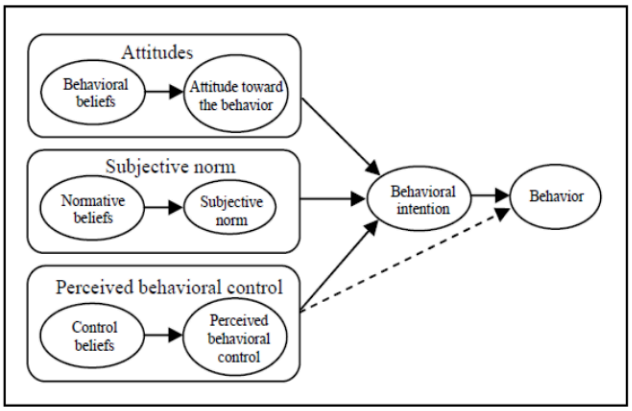

Figure 3 Theory of planned behavior (Piriyawat, 2008)

According to the Theory of Planned Behavior (TPB), the person will behave with causes which they are attitudes toward behavior, deference to Influence of people all around (subjective norm), and their perceptions or beliefs that present their behaviors in the right direction (perceived behavior control). In addition, actual behavioral control does not only believe that it will be controlled, these 3 factors except actual behavioral control still affect to the Behavioral Intention, and aforementioned Intention will push behaving. These 3 factors will be occurred by a fundamental belief as Figure 3.

The last group is a combination (integrated approaches) between first group and second group which is interesting developments because the practical accident cannot be categorized and divided into clear causes. There is a complexity of characteristics and mechanisms in accident occurrences. These developments will enable us to more understand the mechanism of accident occurrences. However, the study still is not found clear principles in integrating factors.

Model Measurements

The study was found that most research was focused on measuring the compatibility of Structural Equation Modeling (SEM) and data collected, determining the influence of direct effect and total effect to explain the causal relationship of latent variables in SEM, and including confirmatory factor analysis of latent variables and observed variables. Nevertheless, the study was found the interesting point of the application of SEM. The research works relating to road accidents still have used correlational research to study the casual relationship of variables which have presented the influence of variables consisting of direct effects and indirect effects. These influences are transmitted through mediator variables or intervening variables.

Figure 4 Mediator and Moderator effects

The study of application of Structural Equation Modeling (SEM) in research works relating to road accident have not found a study of influence transmitted through Moderator variables which can be found in the laboratory research and correlational research(Baron and Kenny, 1986). In the

model, the analysis of one model may have both Mediating effects and Moderating effects in the same model. The characteristics of Mediating effects and Moderating effects have several Mediating and Moderating variables. So, the model is more complex, which conforms to the complexity of accident occurrences.

Summary

Structural equation modeling is becoming widely used in many research fields, such as Marketing, Education, and Psychology. The Confirmatory Factor Analysis (CFA) and the Structural Equation Modeling (SEM) are able to analysis with Observational Variable (MVs), Latent Variable (LVs). Path Analysis (PA) used to analyze and justify type of relationships that is direct or indirect relationship. Therefore, makes it likely that SEM will enjoy widespread used in the future. For SEM application in road accident research should be increase level of analysis to present the influence of variables consisting of direct effects and indirect effects.

References

Anderson, J.C., Gerbing, D.W., 1988. Structural equation modeling in practice: a review and recommended two step approach. Psychological Bulletin 103 (3), 411–423.

Bagozzi, R.P., Yi, Y., 1988. On the evaluation of structural equation models. Journal of the Academy of Marketing Science 16 (1), 74–94.

Baron, R.M. & Kenny, D.A. (1986). The Moderator-Mediator variable distinction in Social Psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51, 1173-1182.

Baumgartner, H., Homburg, C., 1996. Applications of structural equation modeling in marketing and consumer research: a review. International Journal of Research in Marketing 13 (2), 139–161.

Bentler, P.M. and C.P. Chou, 1987. Practical issues in structural modeling. Sociological Methods and Research 16, 78-117.

Bentler, P.M. and K-H Yuan, 1999. Structural equation modeling with small samples: test statistics. Multivariate Behavioral Research 34, 181-197.

Berdoulat, E., Vavassori, D., & Sastre, M. T. M. Driving anger, emotional and instrumental aggressiveness, and impulsiveness in the prediction of aggressive and transgressive driving. Accident Analysis & Prevention(Inpress).

Björklund, G. M. (2008). Driver irritation and aggressive behaviour. Accident Analysis & Prevention, 40(3), 1069-1077.

Bollen, K.A., 1989. Structural Equations with Latent Variables. Wiley, New York.

Bollen, K.A., 1990. Overall fit in covariance structure models: Two types of sample size effects. Psychological Bulletin 107, 256-259.

Boomsma, A., 1982. On the robustness of LISREL against small sample size in factor analysis models. In K.G. Jöreskog and H. Wold (Eds.), Systems under Indirect Observation: Causality, Structure, Prediction, Part 1, 149-173. North-Holland, Amsterdam.

Cantor, D. E., Corsi, T. M., Grimm, C. M., & Özpolat, K. (2010). A driver focused truck crash prediction model. Transportation Research Part E: Logistics and Transportation Review, 46(5), 683-692.

Chin, W.W., Todd, P.A., 1995. On the use, usefulness, and ease of use of structural equation modeling in MIS research: a note of caution. MIS Quarterly 19 (2), 237–246.

Constantinou, E., Panayiotou, G., Konstantinou, N., Loutsiou-Ladd, A., & Kapardis, A. (2011). Risky and aggressive driving in young adults: Personality matters. Accident Analysis & Prevention, 43(4), 1323-1331.

Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med 1997; 126(5):376-80.

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.

Dobson, R., F. Dunbar, C.J. Smith, D. Reibstein and C. Lovelock, 1978. Structural models for the analysis of traveler attitude-behavior relationships. Transportation 7, 351-363.

Donovan, J.E., 1993. Young adult drinking-driving: Behavioral and psychological correlates. Journal of Studies on Alcohol 54, 600-613.

Eboli, l., & Mazzulla, G. (2007). A structural equation model for road accident analysis, 4th international SIIV congress-Palemo. Italy.

Eksler, V. (2010). Measuring and understanding road safety performance at local territorial level. Safety Science, 48(9), 1197-1202.

Fan, X., B. Thompson and L. Wang, 1999. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. Structural Equation Modeling 6, 56-83.

Finch, J.F., S.G. West and D. MacKinnon, 1997. Effects of sample size and nonnormality on the estimation of mediated effects in latent variables models. Structural Equation Modeling 4, 87-107.

Garver, M.S., Mentzer, J.T., 1999. Logistics research methods: employing structural equation modeling to test for construct validity. Journal of Business Logistics 20 (1), 33–57.

Gefen, D., Straub, D.W., Boudreau, M., 2000. Structural equation modeling and regression: guidelines for research practice. Communications of the AIS 1 (7), 1–78.

George, Y., Aggeliki, K., Kallia, A., & Dimitrios, T. (2005). Modelling driver choices towards accident risk reduction. Safety Science, 43(3), 173-186.

Golob, T. F. (2003). Structural Equation Modeling for Travel Behavior Research. Transport research 37B 1–15.

Golob, T.F., and D.A. Hensher, 1996. Driver behavior of long-distance truck drivers: Effects of schedule compliance on drug use and speeding citations. International Journal of Transport Economics 23, 267-301.

Häkkänen, H., & Summala, H. (2001). Fatal traffic accidents among trailer truck drivers and accident causes as viewed by other truck drivers. Accident Analysis & Prevention, 33(2), 187-196.

Hassan, H. M., & Abdel-Aty, M. A. (2011). Analysis of drivers' behavior under reduced visibility conditions using a Structural Equation Modeling approach. Transportation Research Part F: Traffic Psychology and Behaviour, 14(6), 614-625.

Hershberger, S.L., 2003. The growth of structural equation modeling: 1994–2001. Structural Equation Modeling 10 (1), 35–46.

Hoetler, J.W., 1983. The analysis of covariance structures: Goodness-of-fit indices. Sociological Methods and Research 11, 325-344.

Hoogland, J.J. and A. Boomsma, 1998. Robustness studies in Covariance Structure Modeling: An overview and a meta-analysis. Sociological Methods andResearch 26, 329-3.

Iversen, H., & Rundmo, T. (2002). Personality, risky driving and accident involvement among Norwegian drivers. Personality and Individual Differences, 33(8), 1251-1263.

Jöreskog, K.G., 1973. A general method for estimating a linear structural equation system. In A.S. Goldberger and O.D. Duncan (Eds.) Structural Equations Models in the Social Sciences. Academic Press, New York.

Jöreskog, K.G. and D. Sörbom, 1979. Advances in Factor Analysis and Structural Equation Models. Abt Books, Cambridge, MA.

Kaplan, D., 2000. Structural Equation Modeling: Foundations and Extensions. Sage Publications, Newbury Park, CA.

Kline, R.B., (Ed.), 1998. Principles and Practice of Structural Equation Modeling. Guilford Press, New York.

Lee, J.-Y., Chung, J.-H., & Son, B. (2008). Analysis of traffic accident size for Korean highway using structural equation models. Accident Analysis & Prevention, 40(6), 1955-1963.

Loehlin, J.C., 1998. Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis. Lawrence Erlbaum Associates, Mahwah, N.J.

Long, J.S., 1983. Covariance Structure Models: An Introduction to LISREL. Sage, Beverly Hill, CA.

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.

MacCallum, R.C., 1986. Specification searches in covariance structure modeling. Psychological Bulletin 100 (1), 107–120.

Maruyama, G., 1998. Basics of Structural Equation Modeling. Sage Publications, Newbury Park, CA.

Piriyawat, S. 2008. An Application of Norm-Activation Theory and Theory of Planed Behavior for Explaining Intention of Motorcycle-Use Reduction. Research and Development Journal of The Engineering Institute of Thailand. 20 (2), 9–18.

Scott-Parker, B., Watson, B., King, M. J., & Hyde, M. K. A further exploration of sensation seeking propensity, reward sensitivity, depression, anxiety, and the risky behaviour of young novice drivers in a structural equation model. Accident Analysis & Prevention(Inpress).

Scott-Parker, B., Watson, B., King, M. J., & Hyde, M. K. (2012). Confirmatory factor analysis of the Behaviour of Young Novice Drivers Scale (BYNDS). Accident Analysis & Prevention, 49(0), 385-391.

Shah, R., Goldstein, S. M. (2006). Use of structural equation modeling in operations management research: Looking back and forward. Journal of Operations Management, 24(2), 148-169.

Shi, J., Bai, Y., Tao, L., & Atchley, P. (2011). A model of Beijing drivers' scrambling behaviors. Accident Analysis & Prevention, 43(4), 1540-1546.

Shook, C.L., Ketchen, D.J., Hult, G.T.M., Kacmar, K.M., 2004. An assessment of the use of structural equation modeling in strategic management research. Strategic Management Journal 25 (4), 397–404.

Steiger, J., 2001. Driving fast in reverse. Journal of American Statistical Association 96, 331–338.

Sümer, N. (2003). Personality and behavioral predictors of traffic accidents: testing a contextual mediated model. Accident Analysis & Prevention, 35(6), 949-964.

Richer, I., & Bergeron, J. (2012). Differentiating risky and aggressive driving: Further support of the internal validity of the Dula Dangerous Driving Index. Accident Analysis & Prevention, 45(0), 620-627.

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.

Verschuur, W. L. G., & Hurts, K. (2008). Modeling safe and unsafe driving behaviour. Accident Analysis & Prevention, 40(2), 644-656.

Wong, J.-T., Chung, Y.-S., & Huang, S.-H. (2010). Determinants behind young motorcyclists' risky riding behavior. Accident Analysis & Prevention, 42(1), 275-281.

Zhao, L., Wang, X., & Qian, Y. (2012). Analysis of factors that influence hazardous material transportation accidents based on Bayesian networks: A case study in China. Safety Science, 50(4), 1049-1055.