UNDERSTANDING AND IMPROVING DRIVER COMPLIANCE WITH SAFETY SYSTEM FEEDBACK

Mikael Ljung Aust
Sergejs Dombrovskis
Volvo Car Corporation
Sweden
Paper number 13-0455

ABSTRACT

Many crashes usually start with a driver inadvertently leaving the lane. These lane departures broadly fall into two categories. One is where kinematic control is lost, e.g. due to icy roads. The other, and the focus here, is when the vehicle in principle remains controllable, but where the driver for some reason temporarily does not exercise that control. Developing safety systems which detect and act on inadvertent lane departures due to e.g. drowsiness and/or distraction, has a large safety potential.

However, in addition to precise threat detection, successful implementation of such systems also requires an understanding of what motivates and controls the driver's response to system feedback. While threat detection has advanced considerably in recent years, there has yet to emerged a common view on how to understand and improve driver compliance with system feedback in imminent lane departure situations.

The objective of the paper is to formulate a theoretical framework for understanding how safety system feedback is received by the driver in different driving situations. The purpose is to enhance the understanding of what is required to achieve high levels of driver compliance in situations where systems indicate risk, for example of inadvertent lane departures.

The framework is based on the dimensions of perceived threat relevance and opportunities for action. Essentially, when system feedback is received (e.g. a lane departure warning), the driver balances the perceived potential gravity of the situation against the effort required to abide by the system's feedback. This aligns with a general human factors trend toward describing human behavior as a balancing act between goal desirability and energy expenditure.

Application of the framework shows that if the driver associates an imminent lane departure with a low level of threat, correctional effort in response to system feedback will be minimal. To increase lane keeping precision under those circumstances, the vehicle must offer an opportunity for action that requires minimal driver effort to realize. Here, strategies like offering to turn on lane keeping aid as soon as lane keeping starts to degrade might be a way forward. If the driver on the other hand associates a lane departure with a high level of threat, any warning that manages to bring the driver's attention back on the forward roadway will be sufficient. The exception is if the driver is incapable of comprehending or acting the warning, in which case radical actions such as autonomously driving the vehicle to the next rest place might be necessary.

To increase road safety, a deeper understanding of driver compliance is just as important as good threat detection. The issue of how to scientifically approach driver compliance needs to be a top priority in driver behavior analysis. The framework illustrates both the need for, and a viable approach to, a systematic view of how safety system feedback influences driver behavior in lane departure situations. While a step forward, much work remains before the principles governing driver compliance in potentially threatening situations are fully understood.

INTRODUCTION

Active safety systems that detect and act on potentially critical driving situations by alerting or warning the driver to e.g. inadvertent lane departures and imminent lead vehicle collisions have a large safety increasing potential. However, for these systems to accomplish what they set out to do, they must influence driver behavior in the way system designers have intended. For this to happen, it follows that implementation of these systems not only requires precise threat detection and an intuitive driver interface. Also, a deep understanding of what motivates and controls driver compliance with the alerts and warnings given is required.

A concrete example serves to illustrate the point. Say a vehicle manufacturer launches a safety function that with 100 % certainty detects when a driver is close to falling asleep behind the wheel, and alerts the driver to this condition. What happens next? From a system design point of view, the desired response would be that the driver almost immediately stops, drinks a cup of coffee and then takes a fifteen minute nap before continuing the drive.

However, as most can testify from their own experiences as drivers, drivers' real world behavior do not always conform with system design intent. Apart from not always having a thermos of coffee ready for situations like these, drivers normally weight in the system's recommendation together with a number of other factors, such as distance to home or the availability of a safe parking place, before deciding whether to go on or to stop, and it's not obvious that the system will have the final word on the matter.

As the example illustrates, the purpose of in-vehicle alerts and warnings is to guide driver behavior in certain ways, and if this guidance is unsuccessful, so is the system. For safety systems general, a lack of compliance means that the benefit one predicts based on installing the system never will come to fruition.

Given this truly central role of driver compliance in the effect of warning based in-vehicle systems, one would assume that the study of compliance would be a large field. However, there are actually relatively few studies in the empirical warnings literature that assess people's compliance behavior [1].

One of the underlying reasons is fairly straightforward. Compared to studies that focus on pre-cursors to compliance such as warning awareness and comprehension, driver compliance is difficult to study. Coming back to the drowsiness detection function exemplified above, when one starts to think about it, operationalizing how to measure compliance with system output is quite difficult. Is it enough to ask drivers whether they stop in response to warnings given, or are more objective measures required? And if they do stop, how soon after the warning need they stop for it to count as warning compliance and not just stopping because they are tired (which is why they get the warning after all)?

Another reason why compliance is hard to understand and move forward is that the available studies to a certain extent contain results that are less easy to interpret and understand. Some very interesting findings that inspired the current paper come from the recently concluded euroFOT project [2]. In euroFOT, the drivers attitude towards, and usage of, several in-vehicle warning systems were studied. The result which concerns us here is that in this study, two systems that in many aspects are very similar were rated very differently by the users. These systems were Forward Collision Warning (FCW) and Lane Departure Warning (LDW). Both warn the driver when driver action is required, either to avoid leaving the lane or to avoid a conflict with the lead vehicle. The difference between them in euroFOT was that while FCW was highly appreciated by most drivers on many compliance related dimensions such as usefulness and warning relevance, LDW received much less favorable ratings. Moreover, at the same time as LDW was rated as less favorable, the intervention system Lane Keeping Aid (LKA), which countersteers when the driver is about to cross the lane marker rather than sounding a warning like LDW, received very positive feedback from the drivers.

It is far from obvious how these findings should be interpreted. Why do drivers like a warning for longitudinal conflicts but prefer a steering intervention when lateral loss of control is at hand? Why not the opposite? It is clear that a larger picture that describes the more basic mechanisms of what makes drivers adhere or not adhere to warnings is missing. There is simply a lack of agreed-upon concepts and principles when it comes to describing the mechanisms that govern driver compliance with different in-vehicle alerts and warnings.

A CONCEPTUAL FRAMEWORK FOR DRIVER COMPLIANCE

A set of principles and concepts that capture the fundamental ideas of a field of science is often referred to as a conceptual framework. That is what is missing from the research field of behavioral compliance, and also what the current paper tries to address. To exemplify more concretely what is meant by a conceptual framework, it helps to consider work in a related field, i.e. injury prevention. Here, William Haddon formulated a conceptual framework through what can be called the energy transfer model [3]. Haddon stated that injuries occur when "energy is transferred in such ways and amounts, and at such rates, that inanimate or animate structures are damaged" [4].

This basic idea, albeit simple in retrospect, had a tremendous impact on work in injury protection. Accident investigators understood they should collect field data on the ways in which sudden and unwanted energy transfers into humans occur. Those who develop countermeasures realized they should focus creating ways of redistributing unwanted mechanical energy in time and space to avoid it reaching humans. All in all, Haddon's framework presented a simple yet scientifically sound model that practitioners in the field could use as a basis for discussions.

The present paper is an attempt to formulate a conceptual framework for driver compliance with warnings and alerts that can describe the effect of in-vehicle warnings on actual driver behaviour in a way similar to how Haddon's energy transfer model explains the injury preventive function of seat belts, etc. The purpose of the framework is to enhance the understanding of what is required to achieve high levels of driver compliance with system recommendations in potentially critical situations. The framework presented here draws heavily upon the conceptual framework for evaluation of active safety systems presented in [5], though here it is tweaked to suit application towards compliance issues rather than warning design.

While the central concept in Haddon's framework is negative energy transfer, the central concept in the currently proposed framework is control. Control in general can be defined as an ability to direct and manage the development of events [6], or more specifically the maintenance of a goal state in face of disturbances [7]. In the domain of traffic safety, driving can be viewed as a control task that involves continuous adaptation to a changing environment, in a way which promotes goal fulfillment [7] .

Moreover, controlling a vehicle normally involves the pursuit of multiple goals. These can often be described as hierarchically ordered, i.e. a high-level goal can be to reach the destination in time, while lower-level goals include avoiding colliding with lead vehicles and driving within the lane. Such hierarchies of goals is reflected in many driving models [6][8] [9][10].

However, while the hierarchical models above can describe the multiple control processes involved in driving, they cannot account for how the goals, or reference values, are selected. They therefore needs to be complemented by an explanation of why drivers choose the goals they do. One early such account is the zero-risk theory by Näätänen and Summala [11], which proposes that driver behaviour is shaped by excitatory "forces". These forces motivate the driver to actively look for and exploit opportunities for action present in the environment. For example, if a driver wishes to travel faster than the vehicle in front, s/he will look for a gap in the left lane to overtake in.

To keep things balanced however, the excitatory forces are kept in check by inhibitory forces. Originally, Näätänen and Summala [11] proposed that inhibitory forces are driven by subjective risk estimates. Vaa [12] developed that general idea by incorporating Damasio's concept of somatic markers [13]. Vaa states that adaptive driver behaviour largely is governed by physiological reactions to threatening situations, i.e. emotions, experienced by the driver as unpleasant feelings. Somatic markers are emotional signals that attach positive or negative values to opportunities for action and their outcomes. Following Vaa's model, [14] proposes a generalization of the zero–risk model where driver strives to maintain a state of zero discomfort rather than zero risk. In this model, the driver's selection of goals in the driving control processes becomes a balance between a desire for goal fulfillment and discomfort avoidance. This results in adaptive behaviour, with the driver responding to changing driving demands (current and predicted) by seeking reference values which will result in goal fulfillment without generating feelings of discomfort.

In terms of which goals are attractive or not, drivers normally seek goals they believe are within a safety zone [14]. The safety zone is defined as the region of all goal states for which control is maintained. For any other state, non-recoverable loss of control will occur. The important point here is that according to Summala, in order to maintain their state of zero discomfort, drivers generally avoid goal states that are close to the safety zone boundary. Rather, they prefer goals with a certain minimum distance, or safety margin, to the boundary.

The region defined by this safety margin can be conceptualized as a comfort zone, i.e. a region of possible goal states for which no discomfort is felt or predicted by the driver, and which the driver therefore prefers to stay within. As long as the comfort zone boundary is not exceeded, the exact goal state that is chosen does not matter very much. However, if the comfort zone boundary is exceeded, a feeling of discomfort will be experienced, and the driver will take adaptive action to reduce that feeling.

These concepts are illustrated in Figure 1 below. The comfort zone contains the speeds for which the driver expects no feeling of discomfort, given the subjective assessment of road friction. The safety margin is the difference between the comfort zone boundary and the safety zone boundary (i.e. between the maximum speed that the driver is comfortable with and the maximum possible speed which does not lead to skidding). In this example, the driver successfully perceives the change in safety zone boundary which occurs when friction is reduced due to for example a sudden snowfall. Since the current speed feels uncomfortable in relation to this change in conditions, the driver adapts by slowing down to a speed well below the safety zone boundary for the new friction conditions, and manages to do so without exiting the comfort zone. The driver thus avoids feelings of discomfort as well as loss of control.

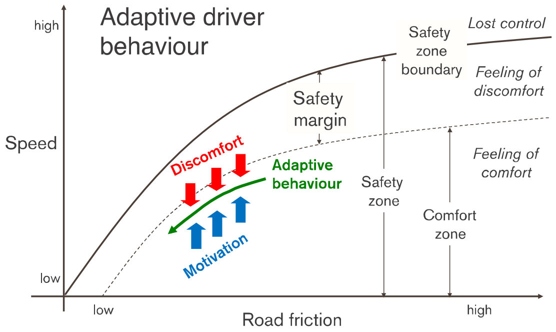

Figure 1: Driver adapting speed when road friction changes

As stated in the introduction, the purpose of a conceptual framework is to capture the fundamental ideas of a field of science. The authors believe that the conceptual framework presented here does exactly that for the field of driver compliance with in-vehicle warnings and alerts, based on the following reasoning.

The design intent of all in-vehicle warnings is to tell the driver that a certain safety margin threshold has been exceeded. The exceedance in itself can be related to many different measures of safety, such as acceptable levels of drowsiness, lateral positioning or kinematic margins to the lead vehicle. Overall however, the one thing all warning and alert systems have in common is that their designers think some safety margin, however defined, should be larger than what it currently is.

Put in terms of the conceptual framework described above, an alert or a warning is principally a sign that the driver has exceeded the system designer's comfort zone boundary, and the designer therefore wants the driver to take corrective action in order to increase the safety margin. From this follows that a key enabler for high levels of driver compliance with alerts and warnings is that the system designers and the driver's view of the situation match, i.e. that they share the same definition of where the comfort zone boundary is. If they do, then all is well. The driver when alerted will realize the safety margin is getting uncomfortably small and will adjust accordingly. If they do not however, the driver will regard the system's output as a nuisance and general source of irritation.
In short, driver compliance crucially depends on how well the system designer's assessment of the comfort zone boundary matches the driver's assessment. This is the key idea of the conceptual framework proposed in this paper.

HOW TO UNDERSTAND WHERE THE DRIVER'S COMFORT ZONE BOUNDARY LIES

In view of the framework, a key research topic for systems designers is to find out how drivers' comfort zone boundaries are set and altered. Here, some key dimensions can be identified.

Expectations about consequences

One very important determinant seems to be the drivers' expectations. Expectations are the anticipatory outcome of a behaviour [15] and are comprised of a person's attitudes and beliefs. The expectations that a person brings to a situation influences whether s/he notices a warning, how s/he processes the warning stimuli, and whether s/he complies. A person's knowledge base, whether accurate or not, will therefore influence behaviors coupled to that knowledge.

In a general context, it is also well known that warning effectiveness tends to increase as a function of perceived hazardousness [16]. Hazard perceptions in turn are generally believed to be based on both the perceived likelihood (probability of experiencing an undesirable outcome) and severity (seriousness of the consequence) of potential incidents [17]. Of these two, severity of injury seems the better predictor of risk perception [18][19]. In other words, people will be more wary of rare events that can do a lot of harm than of more frequent events that do less harm.

This has important implications for warning system design. If the system is able to convey that severe bodily harm is a potential consequence of going beyond the comfort zone boundary as defined by the system designers (even though that outcome may be unlikely), the driver is more likely to comply with the warning. If not, the driver will be less likely to comply, since the current driving state will be perceived as being within the comfort zone boundary.

Familiarity and warnings – the risk of crying wolf too often

In studies investigating the relationship between familiarity and perceived risk, results indicate that increased familiarity with a product reduces it's perceived hazardousness [20]. People may become desensitized to warnings as a result of repeated exposure without immediate consequences [21]. Or put differently, when benign experiences occur, they affect the expectation of risk.

For in-vehicle warning design, the consequence is that repeated warnings in benign situations, i.e.

where nothing bad happens, will lead to a decrease in the perceived hazardousness of the situation. In other words, the further down the timeline of warning exposure that the driver is, the more perceivably hazardous the situation actually has to be in order for the driver to respond as intended. Every benign warning gives the driver reason to believe that the warning is given within the comfort zone, and hence there is no need to regard it as relevant for driving.

Another consequence is that since systems normally are improved over time, applying a conservative warning strategy when releasing new systems is warranted. It seems better to warn rarely at high levels of threat compared to warning more often at lower levels of threat, because if the first warnings are not perceived as relevant, later warnings will not be either. Familiarity will step by step push the warning further inside the comfort zone, unless some transformative (i.e. near crash or crash) experience intervenes.

Illusory superiority and compliance

Another important determinant of the comfort zone boundary is what often is referred to as threat denial. Sometimes people respond to hazard warnings with feelings of personal immunity or overconfidence [22][23].

For example, Svenson [24] surveyed 161 students in Sweden and the United States, asking them to compare their driving safety and skill to the other people in the experiment. For driving skill, about 9% of the US sample and 70 % of the Swedish sample put themselves above the median. For safety, 88% of the US group and 77% of the Swedish sample put themselves in the top half. In a similar study, McCormick, Walkey and Green [25] asked 178 participants to evaluate their position on eight different driving skill dimensions. Only a few rated themselves as below average at any point. When all dimensions were considered together, about 80% of the participants evaluated themselves as being better than the average driver.

For in-vehicle warnings, the implication is that the warning has to be designed to break through what can be called the overconfidence boundary. This means that a warning cannot be presented in a way that primarily appeals to the average likelihood of risk for example, because only average drivers are susceptible to average risk, and oneself is by definition a better driver than average. Instead, for a warning to be perceived as relevant, it has to trigger somatic markers in the driver, making him/her perceive the situation as immediately uncomfortable and requiring corrective action.

INTERPRETING THE EUROFOT RESULTS IN LIGHT OF THE FRAMEWORK

Based on the framework, it can be assumed that when the driver regards system output as irrelevant (i.e. not associated with a threat that can result in serious injury), the correctional effort made by the driver will be minimal. This helps explain the findings from euroFOT mentioned in the introduction [2]. Translated into the framework, the forward collision warning is appreciated because at the point in time when the warning is given, the driver agrees that it could potentially lead to severe bodily harm if s/he does not brake (i.e. a collision with the lead vehicle is close at hand). However, the sense of potential bodily threat when crossing the lane marker is much lower since there normally is a certain portion of road shoulder without apparent obstacles available.

Drivers thus literally do not feel the need for LDWs, because there is no somatic marker that elicits discomfort in the driver if s/he chooses not to act when crossing the lane marker. In the FCW case on the other hand, the possibility of crashing into the lead vehicle does trigger a somatic discomfort response. Drivers thus appreciate the warning and are likely to act on it, or if not, at least agree that it is relevant.

This presents an interesting conundrum for system designers that wish to decrease the number of crashes that are initiated by an inadvertent lane departure. From an engineering standpoint, the chain of events that lead to these crashes start when the vehicle leaves the lane, and logically speaking, the lane marker should thus be equivalent to the comfort zone boundary. However, drivers apparently view things differently. They seem to treat the lane marker more as a useful recommendation about where to drive rather than as an unbreachable boundary. In their minds, they are not afraid of lane departures, they are afraid of road departures. Therefore they show much bigger respect for other vehicles than for lane markers.

System designers therefore probably need to reconsider their approach to the problem of crashes that start with inadvertent lane departures. Since drivers generally are not afraid of the potential consequences of a lane departure, one approach would be to offer an opportunity for action in terms of staying in the lane that requires a minimum of effort to carry out. Automatically steering the vehicle back when an imminent lane departure is detected might therefore be a way forward, and this is indeed what the driver feedback on the LKA system in euroFOT shows [2]. Another approach would be to modify the warning strategy. For example, if the warning comes close in time to the vehicle leaving the road rather than the lane, or if one warns only when there is an oncoming vehicle in the lane you're drifting into. In these cases, the driver's threat assessment is more likely to correspond with the warning, and a higher level of compliance can be expected. To illustrate the point, think about whether you would appreciate a lane departure warning when there is six feet of shoulder followed by a guardrail, compared to when there is a 500 feet drop one foot outside the lane marker and no guardrail.

DISCUSSION AND LIMITATIONS

Application of the framework shows that if the driver associates an imminent lane departure with a low level of threat, correctional effort in response to system feedback will be minimal. To increase e.g. lane keeping precision under those circumstances, the vehicle must offer an opportunity for action that requires minimal driver effort to realize. Here, strategies like offering to turn on lane keeping aid as soon as lane keeping starts to degrade might be a way forward. If the driver on the other hand associates a lane departure with a high level of threat, any warning that manages to bring the driver's attention back on the forward roadway will be sufficient.

To increase road safety, a deeper understanding of driver compliance is just as important as good threat detection, and the issue of how to scientifically approach driver compliance needs to be a top priority in driver behavior analysis. The framework presented here illustrates both the need for, and a viable approach to, a systematic view of how safety system feedback influences driver behavior in lane departure situations. While this is a step forward, there obviously remains a lot of work before all principles that govern driver compliance in potentially threatening situations are fully understood and accounted for.

A limitation of the current framework is that is driver and vehicle centered. There are however certain situations where a high degree of in-vehicle compliance only is the first step toward a good solution. For example, even if the driver intends to comply with an alert from a drowsiness warning system, it may sometimes simply not be possible to do so. The nearest highway exit may be miles away, or rest places that feel safe might be scarce. Certain enablers for compliance thus exist outside the vehicle, and this must not be forgotten in the process.

REFERENCES

[1] Kalsher, M.J, and Williams, K.J (2006) "Behavioral compliance: theory, methodology and results" in Wogalter, M.S. (ed.) *Handbook of warnings*. Mahwah, N.J.: Lawrence Erlbaum.
[2] Sanchez, D., Garcia, E., Saez, M., Benmimoun, M.,, Pütz, A., Ljung Aust, M., Gustafsson, D., Metz, B., Saint Pierre, G., Tattegrain, H., Guidotti, L., Schindhelm, R., Heinig, I., Malta, L., Obojski, M.A. (2012). "Final results: User acceptance and user-related aspects" Deliverable 6.3 of the European Large-Scale Field Operational Test on Active Safety Systems (euroFOT) project. Brussels: European Commission.
[3]Haddon, W. (1968). "Changing approach to epidemiology prevention and amelioration of trauma - transition to approaches etiologically rather than descriptively based". *American Journal of Public Health and the Nations Health,* 58(8), 1431-1438.
[4] Haddon, W. (1973). "Energy damage and 10 countermeasure strategies". *Journal of Trauma-Injury Infection and Critical Care,* 13(4), 321-331.
[5] Ljung Aust, M., & Engström, J. (2011). "A conceptual framework for requirement specification and evaluation of active safety functions". Theoretical Issues in Ergonomics Science, 12(1), 44 - 65.
[6] Hollnagel, E., & Woods, D. D. (2005). "Joint cognitive systems: foundations of cognitive systems engineering". Boca Raton, FL: CRC Press.
[7] Engström, J., & Hollnagel, E. (2007). " General Conceptual Framework for Modelling Behavioural Effects of Driver Support Functions". In P. C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments Critical Issues in Driver Interactions with Intelligent Transport Systems.* London: Springer. pp. 61-84.
[8] Allen, T.M., Lunenfeld, H., & Alexander, G. J. (1971) "Driver information needs". Highway Research Record, 366(366), 102-115.
[9]Michon, J.A. (1985). "Critical view of driver behavior models: what do we know, what should we do?" In Evans, L. and Schwing, R. (eds) *Human Behaviors and Traffic Safety*, Plenum Press, New York
[10] Hollnagel, E. (2003). "Handbook of cognitive task design". Mahwah, N.J.: Lawrence Erlbaum.
[11] Näätänen, R., & Summala, H. 1976. "Road-user behavior and traffic accidents". Amsterdam: New York.
[12] Vaa, T. (2007). "Modelling Driver Behaviour on Basis of Emotions and Feelings: Intelligent Transport Systems and Behavioural Adaptations". In P.C. Cacciabue (ed). *Modelling driver behaviour in automotive environments: critical issues in driver interactions with intelligent transport systems* London: Springer. pp. 208-232.
[13] Damasio, A.R. (1994) Descartes' error: emotion, reason, and the human brain. New York: G.P. Putnam.
[14] Summala, H. (2007) "Towards Understanding Motivational and Emotional Factors in Driver

Behaviour: Comfort Through Satisficing". In P.C. Cacciabue (ed). *Modelling driver behaviour in automotive environments: critical issues in driver interactions with intelligent transport systems.* London: Springer. pp. 189-207.
[15] Bandura, A., (1986) "Social foundations of thought and action: a social cognitive theory". Englewood Cliffs, NJ: Prentice- Hall.
[16] DeJoy, D.M. (1991) "A revised model of the warnings process derived from value-expecancy theory". Proceedings of the Human Factors Society, 35, 1039-1042.
[17] Slovic, P., Fischhoff, B. and Lichtenstein, S (1979) "Rating the risks". *Environment, 21*, 14-20, 36-39.
[18]Wogalter, M.S., Brelsford, J.W., Desaulniers D.R., and Laughery, K.R (1991) "Consumer product warnings: the role of hazard perception". Journal of Safety Research, 22, 71-82.
[19] Young, S.L., Brelsford J.W., and Wogalter, M.S. (1990) "Judgments of hazard, risk and danger. Do they differ?" Proceedings of the Human Factors Society, 34, 503-507.
[20] Vredenburgh, A.G., and and Zackowitz, I.B. (2006) "Expectations" in Wogalter, M.S. (ed.) *Handbook of warnings*. Mahwah, N.J.: Lawrence Erlbaum.
[21] Stewart, D. W. and Martin, I. M. (1994) "Intended and unintended consequences of warning messages: A review and synthesis of empirical research". Journal of Public Policy & Marketing, 13, 1-19.
[22] Fischhoff, B., Slovic P. & Lichtenstein, S (1977) "Knowing with certainty: The appropriateness of extreme confidence". Journal of Experimental Psychology: Human Perception and Performance, 3: 552-564.
[23] Slovic, P., Fishhoff, B., & Lichtenstein, S. (1980). "Facts and Fears: Understanding Perceived Risk". In Albers, W.A. (Ed.) *Societal Risk Assessment: How Safe Is Safe Enough?* New York, NY: Plenum Press
[24] Svenson (1981) "Are we all less risky and more skillful than our fellow drivers?" Acta Psychologica, 47, 143-148, North-Holland Publishing Company.
[25] McCormick, I.A., Walkey, F.H. & Green, D.E. (1986). "Comparative perceptions of driver ability – A confirmation and expansion". Accident Analysis & Prevention, 18, 205-208.