Trust and Traffic Safety

Ola Boström

This lecture will discuss the critical role of trust as it relates to traffic safety and achieving the vision of zero road fatalities and severe injuries.

My career path started in theoretical physics before I got the chance to do research in biomechanics. My focus for many years was to understand injury mechanisms during car crashes. With that understanding it was possible to come up with and try out new concepts for injury protection, such as anti-whiplash seat designs, far-side airbags, and seat belts for obese occupants. However, to be able to evaluate new concepts of protection, in most cases the first and most time-consuming step was to come up with new tools like injury criteria [1-3]. I learned early the importance of global collaboration, trust and governmental funding. At the time I changed my career path from physics to biomechanics, whiplash injury was ill-defined, ill-diagnosed, disbelieved, and still the number one priority for many insurance companies. Among the researchers I networked with (from Europe, US and Canada), we not only trusted each other and shared what we found in our analyses and experiments, we trusted the stories of the victims. That trust, combined with the backing of governmental funding, was probably the main reason we succeeded in establishing the necessary tools (e.g., crash test dummy, injury criteria, standardised crash pulse, etc.) for the development of new vehicle seat designs and for the tools later to be globally accepted and used. To the delight of the Swedish government (having funded through Vinnova) Volvo cars and Saab were the first to market with seats that mitigated whiplash injury.

Working for Autoliv (the only company I ever applied to work for) felt good, knowing our airbags and seat belts saved 25 000 lives each year. However, our CEO later challenged us to prevent many more fatalities and to reduce the consequences of other injuries. By 2013, I was Vice President of Research and realised that although biomechanics was still the foundation of traffic safety, the field of human factors or behavior must also be addressed to successfully tackle the more than million lives lost each year on the roads [4]. Note that half of these victims were unprotected outside the vehicles. Passive safety in the form of airbags and seat belts is a prerequisite for active safety, but of what use is active safety in the form of driver assistance (e.g., lane keeping assist, cruise control, etc.) or emergency actions (e.g., automatic emergency brakes, etc.) when they can be switched off, overridden, or only partially trusted.

So, let's dive into trust—the main topic of my lecture—and more specifically the trust between humans and humans as well as the trust between humans and machines. Imagine you are traveling in a vehicle on an undivided road and you encounter an oncoming vehicle in the opposite lane. If you and the driver in the approaching vehicle are both in control of your vehicles, you need to trust yourself and the other driver to remain comfortable enough to keep driving. If you or the other driver have delegated some of the driving tasks to an algorithm, you need to trust the vehicle technology and its manufacturer. Clearly trusting people is different from trusting technology. My boss and I were once invited to a demo by a cyber safety start-up. I was in the back seat and my boss was in the driver's seat. We were driving along a back street and to this day I have never seen such a surprised face when suddenly the brakes completely failed in a staged cyber-attack. We are used to trusting the steering and braking systems (which we believe are mechanical systems), but we are not so confident in computers that need rebooting once in a while. Why then do people trust their fellow drivers sharing the roads and the vehicle technology so far? Beside the fact that most people have not experienced their vehicle being cyber attacked, my answer to this question would simply be the trust that has been created by the biomechanical, crashworthiness and epidemiology research over the years. More about that later.

Today O. Boström is a self employed consultant. His email address is ola.bostrom@gmail.com.

Trust is normally defined as the confidence in or dependence on a person or quality. Humans have the ability and experience of trusting one another, although trust is not always easily gained. Actually, there is an equation for trust, and I believe this equation can be useful for understanding vehicle safety. The trust equation (see below) states that trustworthiness is equal to the sum of credibility, reliability, and intimacy, all divided by self orientation [5].

Trustworthiness = (credibility + reliability + intimacy) / self orientation

Credibility refers to your (intellectual) authority, reliability refers to keeping promises (you do what you say you can do), and intimacy refers to the strength of the relation between the two parties. But the trust created by the factors in the numerator can be ruined by the denominator of self orientation, which refers to self-interest and the lack of transparency (not sharing information). Moreover, and important for this lecture, humans can accept failures in trust depending on the outcome. You do trust a plumber, for example, if they are licensed (credible), known to keep a timeline (reliable), and recommended by friends you already trust (intimacy). Even if there is a leak after the plumber has completed their work, you may keep trusting the plumber if they are transparent about why the leak happened (self orientation).

Most of us have experienced a vehicle crash (including a fender bender) and still continue to drive and trust other drivers (and perhaps the algorithms). I believe the main reason for this forgiveness is the nature of our passive safety systems, which adopt biomechanical limits and allow for failures in human behaviors on the road (Vision Zero safety philosophy). Are we then willing to repeatedly forgive a technology and is it then adequate or valid to apply the same trust equation to a machine or algorithm? Probably not, but it could be a useful starting point.

There are some alarming indications that some drivers develop a false or inflated sense of trust in vehicle technology. This “overtrust” can lead to drivers giving critical driving tasks to a poor algorithm based on poor sensor data. Just as it takes time to develop trust in another human, it must take time to develop trust in a machine, particularly when the machine is in control of life-critical tasks. It is not a good idea to perform experiments with drivers when the consequences are life-threatening. Nevertheless, if trusting machines to do what they are designed to do is the next big step toward Vision Zero, let’s continue our quest (refraining from life-threatening experiments). Ask yourself, do you trust the blind-spot detection system to make a lane change without looking over your shoulder? Hopefully not, because this Advanced Driver Assistance System (ADAS) function is not meant to be trusted in this way. Yet drivers become dependent on this system and tell stories when they were saved from a crash by the sudden blink in the side mirror. Do you trust the automatic emergency brake (AEB) system to detect and stop for pedestrians? Hopefully not, the performance of AEB systems in nighttime is often poor, as shown recently by the Insurance Institute for Highway Safety (IIHS) [6]. While some basic-rated vehicles in the IIHS tests avoided hitting the pedestrian dummy at the lower test speeds, none of the systems avoided collisions in the higher-speed crossing or higher-speed parallel scenario.

During my career I was intrigued and inspired by the CEO I was working for when he asked for a blind-spot detection you could trust to make a lane change. For me this challenge worked as a vision to develop an exciting new technology. This challenge was easy for me to relate to and it reminded me of when I was younger and used to make a left turn in a crossing without looking to the right, solely depending on my friend (who I trusted) in the passenger seat who had just said “OK right!”. I really would like to be able to trust a machine as much as I trusted my friend. A more realistic possibility in the near future is an AEB system that you can trust in most weather environmental conditions to detect and act in time.

When it comes to the denominator in the trust equation, a human would not trust a machine if they suspected the machine to have its own self-interest. And that is why transparency is so important to creating trust. The machine (or machine designers/manufacturers) must be very clear on its intended mode and priorities of operation. Also, this trust works both ways: for the machine to take critical tasks (but not all) from the driver, the

machine needs to trust the human. This is why driver monitoring is so important, a technology now being globally regulated and rated.

In biomechanics we may continue to shift our focus from life-threatening injuries towards a wider range of injuries and continue expand the range of loading conditions, including seating postures and sudden AEB activation. Remember that Vision Zero doesn't mean zero crashes, and the closer we come to achieving Vision Zero, the more the remaining injuries will be a consequence of an AEB system having mitigated the crash. In human factors, we need to understand how to design technology that supports real trust between humans and machines. Regarding active safety technology, AEB needs to be evaluated and rated under complex scenarios (credible), shown to work under real conditions (reliable), be state-of-art (intimacy), and a clear statement of the Operational Design Domain (operating context defined by a set of conditions, including environmental, geographical, time of day) sharing its strengths and shortcomings (self-orientation).

Back to my previous statement, I strongly believe we have the level of passive safety in today's vehicles thanks to basic research, not necessarily ground-breaking or applied research, performed over the years in the fields of biomechanics, epidemiology and crashworthiness. The end-consumers, the rating and regulation organizations, and the vehicle manufacturers—all of whom are responsible for the safety of the vehicle fleet of today—are humans after all. And all of these people, without necessarily thinking about it, have probably relied on the trust equation.

We may take the safety of the vehicle fleet for granted, but think about it, the IRCOBI community would have a top score when it comes to trust. The IRCOBI community is one of a few global gatherings (credible), has annual conference (reliable), constitutes a solid network open for new students (intimacy), and acts out of curiosity and safety compassion (self-orientation). Let's take the next steps, continuing to develop our understanding of how humans are injured (body) and fail to comply (mind), as well as demanding transparent AEB systems for vehicles that function day and night, rain or sun.

I. REFERENCES

1. Bostrom 0, Svensson MY, Aldman B, Hansson HA, Häland Y, Lövsund P, Seeman T, Suneson A, Säljö A, Örtengren T (1996) A new neck injury criterion candidate - based on injury findings in the cervical spinal ganglia after experimental neck extension trauma. Proc 1 996 Int Conf on the Biomechanics of Impact (IR CO BI), Dublin, Ireland 1 23-136
2. Bostrom O, Gabler H, Digges K, Fildes B, and Sunnevang C, Injury Reduction Opportunities of Far side Impact Countermeasures Ann Adv Automot Med. 2008; 52: 289–300.
3. Forman J, Lopez-Valdes F, Lessley D, Kindig M, Kent R, Bostrom O, The effect of obesity on the restraint of automobile occupants, Ann Adv Automot Med. 2009 Oct:53:25-40.
4. Maister D, Galford R, Green C, 2000, The Trusted Advisor, ISBN 9780743207768
5. Road traffic injuries, 2003, WHO news
6. Autobrake performance slips after sundown, 2022, IIHS news Aug 30