

The dilemma of autonomous driving: Reflections on the moral and legal treatment of automatic collision avoidance systems

Eric Hilgendorf*

Introduction

The fact that technological progress constantly raises new legal problems is already almost a platitude. Remarkably, however, it seems occasionally to lead to old legal problems reappearing under new guises. A much discussed example currently is the problem of how algorithm-controlled collision avoidance systems, as are used, for example, in modern automobiles, cause their vehicles to react in life-threatening emergency situations. Suppose a vehicle equipped with such a system approaches an accident scene. Three severely injured accident victims, A, B and C lay unconscious on the road, but victim D was able to drag himself to the side of the road and is grasping a sign post to stay on his feet. A second vehicle is approaching the scene. It is moving too fast to stop. It is also not possible for it to swerve in such a way as to avoid striking A, B, C and D. How should the onboard computer steer the car? The attractiveness of such hypothetical cases is due not least to the fact that they help to illustrate the basic values of a legal culture, in a manner which is also accessible to a broader public. Proposed solutions sometimes take on the character of legal and social policy decisions¹.

In this contribution, a proposal to deal with the above mentioned problem will be developed that meets practical requirements, but at the same time is consistent with German legal doctrine. Towards this end, the no-

* Prof. Dr. Dr. Eric Hilgendorf, Julius-Maximilians-Universität Würzburg.

1 In the current debate, the decision-making problem sketched out above is often reduced to the opposition “Kant vs. Bentham”, especially in more popular representations, which, however, certainly remains inadequate because of the chauvinistic Germanic undertones of many such comparisons. For the political-historical dimension of the distinction between “German culture” (Kant), “shallow” French “civilization” (Voltaire) and the “utilitarian merchant spirit” of the British (Bentham), see Hilgendorf, “Rechtsphilosophie der Aufklärung” in Hilgendorf & Joerden (eds.), *Handbuch der Rechtsphilosophie*, 2017, p. 137 et seq.

tion of *degrees of wrong* will be introduced and the already established legal concept of *accepted risk* will be extended to automatic technological systems.

I. Automated driving and the law

The ethical and legal issues that have arisen in the context of automatic collision avoidance systems in motor vehicles have become an important issue in the debate on the future of road transport in Germany². The new possibilities provided by automated driving are should definitely be rated positively overall – one need only consider benefits such as mobility gains for the elderly and the disabled, improvements in road safety, environmental protection, energy efficiency and an increase in the ease of transport³. It would therefore be wrong to view automated driving from the outset with skepticism or to reject it. The law should not block, but rather should steer and promote the development of important new technologies; According to the view represented here, technology law should therefore not be an instrument for preventing innovation, but rather for supporting (and promoting) innovation.

It is obvious, however, that automated driving poses a multitude of difficult and unresolved legal problems. This applies to international law (in particular the Vienna Convention on Road Transport, 1968), as well as national constitutional law, civil liability law, criminal law, data protection law, technical approval law, and insurance law⁴. In this article I will try to analyze a particularly controversial problem at the interface between

-
- 2 The starting point of the current debate were articles by the American philosopher of technology Patrick Lin, cf. for example Lin, *The Ethics of Saving Lives with Autonomous Cars is Far Murkier than you Think* (<https://www.wired.com/2013/07/the-surprising-ethics-of-robot-cars/>); Lin, “Why Ethics Matters for Autonomous Cars” in Maurer *et al.* (eds.), *Autonomes Fahren*, 2015, pp. 69 – 85. Cf. also Bonnefon, Shariff & Rahwan, *Autonomous Vehicles Need Experimental Ethics: Are We Ready For Utilitarian Cars?* (https://www.researchgate.net/publication/282843902_Autonomous_Vehicles_Need_Experimental_Ethics_Are_We_Ready_for_Utilitarian_Cars); dies., *The Social Dilemma of Autonomous Vehicles* (Science on 24 Jun 2016: Vol. 35, DOI: 10.1126/science.aaf2654).
 - 3 Hilgendorf, “Gutachten zum Thema ‘Automatisiertes Fahren und Recht’” in 53. *Deutscher Verkehrsgerichtstag 2015*, 2015, pp. 55 – 72 (57 *et seq.*).
 - 4 Cf. also the overview in Hilgendorf, *op.cit.* 2015 (Fn. 3). p. 59 *et seq.*

ethics, constitutional law, criminal law and civil law, namely the question of which rules may or should be incorporated into automatic collision avoidance systems⁵. What we are dealing with here are systems of rules that enable the on-board computer of a vehicle to avoid an obstacle in the direction of travel, and this is done by the vehicle much faster than would be possible for a human being, who, in such a situation, can neither assimilate the necessary information quickly enough nor turn in time to change the trajectory of the vehicle to avoid a collision.

The new collision avoidance systems are likely to contribute to a reduction in the number of road traffic accidents. They will, however, also cause accidents if swerving vehicles are steered towards targets that would not have been hit in the absence of the computer directed evasive maneuvers. In this respect, the situation is similar to what happened when airbags or seatbelts were first introduced. That was also highly controversial at the time, since the devices do not merely save lives and prevent injuries, but in a small number of individual cases can cause injuries or deaths⁶.

Of course, collisions occur in road traffic today, including those involving injuries or even deaths. Human car drivers are frequently overwhelmed in collision situations and can then no longer make well thought out decisions. This is also one reason why collision scenarios in road traffic have so far hardly been analyzed either from ethical or legal perspectives. The new possibilities offered by technology compel us to consider and analyze the relevant processes and sequences of events. One could even say that a *compulsion to analyze and to explicate* exists in association with the development of algorithms, which, in parallel with the introduction of new autonomous systems, is impacting the way we live and work. Sequences of events that previously were more or less uncontrolled, and indeed unfolded in an uncontrolled manner, can now be decompiled into individual elements and processed in a structured way using algorithms. They can then be steered and controlled.

As (causative) factors relevant to collisions become more transparent and more controllable, responsibility arises, namely both in moral and le-

5 The entirety of implemented rules constitutes a system of norms which for human beings could be characterized as a “fundamental moral orientation”.

6 Bergmann, “Die Gurtdebatte der 1970er und 1980er Jahre in der BRD” in *Technikgeschichte* vol. 76 (2009), pp. 105 – 130; cf. also Forschungsgemeinschaft Der Mensch im Verkehr (ed.), *Für und Wider Sicherheitsgurte*, 1973. See also <http://www.spiegel.de/einestages/einfuehrung-der-gurtpflicht-a-946925.html>.

gal terms. It cannot be avoided by refusing to use automatic collision avoidance systems entirely or in certain accident situations, programming the computer to make random decisions – the decision not to make a decision is also a decision which creates responsibility⁷.

In emergency situations in which an actor cannot avoid violating one (of at least two) legal interests, the *principle of the lesser evil* applies fundamentally in our legal system⁸: damage caused must be kept as low as possible. If the killing of one or more human beings can only be avoided by damaging someone's property, the property damage is justified. The same rule applies in the hypothetical road traffic dilemma discussed above: if a car is involved in a potential accident situation in which it is about to run over several seriously injured people lying on the road, it is imperative that the car swerve to avoid hitting those people even if it then, for example, causes property damage to a sign post, or to a parked car, or to objects standing at the side of the road. The value of the damaged chattels pales in significance – human lives are always more important than things according to hierarchy of values of our legal system⁹.

The principle of the lesser evil becomes problematic, however, when the life of one human being is pitted against the physical integrity or even the life of another human being. What we are confronting here is a fundamental legal and ethical problem involving collision avoidance systems. How should the system decide when one life is pitted against another? Who should live and who should die? There are a range of similar hypo-

7 The concept of responsibility used here can be visualized in the following way: Person X is responsible for an event Z under rule Y. If one accepts this, then among other things it becomes clear that only persons can be responsible for violating a rule. The (socially determined) consequences of an attribution of responsibility can be manifold; their most important manifestations in the law are civil liability (a duty to pay compensation for damage incurred) and criminal liability, that is, the commission of all the elements of a criminal offence, so that conviction and punishment may follow.

8 The most important expression of this principle in German law are the rules governing necessity in the criminal law (§ 34 StGB), under which the protected interest must “significantly outweigh” the interest interfered with.

9 Moreover, there is much to be said for not merely classifying chattels according to their monetary value, but rather also taking into account other considerations, for example with respect of works of art (Michelangelo's Pietà) or animals (we have all heard that some people love their pet more than any human being). In both cases, there are already laws (for example, the Copyright Act for Art and Photography, the Animal Protection Act) which distinguish these chattels from other chattels.

thetical situations discussed in the ethics literature, each with its own name, such as “the plank of Carneades¹⁰”, “castaways on the high seas”¹¹, “euthanasia of the mentally ill during the Third Reich”¹², “the switchman’s case”¹³ and “the trolley problem”¹⁴. Of particular practical relevance in this context is the decision of the German Federal Constitutional Court (2006) on the Aviation Security Act (*Luftsicherheitsgesetz*)¹⁵. At issue was the question of whether a commercial airliner filled with innocent passengers¹⁶, which had been hijacked by terrorists with the intent of using it as a weapon of mass destruction, for example, by crashing it into a

-
- 10 On this subject cf. Hilgendorf, “Tragische Fälle. Extremsituationen und strafrechtlicher Notstand” in Blaschke *et al.* (eds.), *Sicherheit statt Freiheit? Staatliche Handlungsspielräume in extremen Gefährdungslagen*, 2005, S. 107 *et seq.*
- 11 Mitsch, “‘Nantucket Sleighride’ – Der Tod des Matrosen Owen Coffin” in Heinrich *et al.* (eds.), *Festschrift für Ulrich Weber*, 2004, p. 49 *et seq.*; Simpson, *Cannibalism and the Common Law*, 1984; Ziemann, *Zeitschrift für international Strafrechtsdogmatik* 2014, p. 479 *et seq.*
- 12 OGHSt 1, 321; BGH, *Neue Juristische Wochenzeitschrift* 1953, p. 513.
- 13 A railway car rolls down a sloping section of track toward a group of five railroad workers. A switchman can save the lives of the five railroad workers only by redirecting the car onto a side rail where there is a person standing, who will be struck and killed by the car. Can the switchman lawfully redirect the train? This problem, which has been discussed in many different variations, can be traced back to Welzel, *Zeitschrift für das Strafrechtswissenschaft* 63 (1951), p. 47 (51). But switchman cases may also be found in older criminal law writings, e.g. by Köhler, *Der Notstand im künftigen Strafrecht*, 1926, p. 45 comment 1.
- 14 In 1967 the British moral philosopher, Philippa Foot, discussed the switchman problem in her article “The Problem of Abortion and the Doctrine of the Double Effect”, *Oxford Review* 5 (1967) pp. 5 – 15. Since then, what is known by the name “trolley problem”, has been a core element of Anglo-American moral philosophy, most recently, for example, in Edmonds, *Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us About Right and Wrong*, 2014; Kamm, *The Trolley Problem Mysteries*, edited and introduced by Eric Rakowski, 2016; Cathcart, *The Trolley Problem or Would You Throw the Fat Guy Off the Bridge? A Philosophical Conundrum*, 2013.
- 15 BVerfGE 115, 118 *et seq.*
- 16 Below, the terms “innocent” or “innocent person” are used as non-technical descriptions for the designation of two classes of persons. Firstly those who are free of moral wrong or are not culpable so that there is no reason why they should be subjected to or bear special risks, for example of being injured or killed, as would be the case in the course of being judicially sanctioned or punished. Secondly, a class of persons who have not assumed special risk of injury in the performance of their professional duties, for example soldiers, police, fireman, etc.

city centre, could be shot down. In its decision, the court rejected the idea that the airliner could be shot down. The court primarily reasoned that such a course of action would violate the human dignity of the aircraft's passengers. It held that it was unconstitutional for human lives to simply be "weighed against one another"¹⁷.

The dilemma of sacrificing lives in order to save other lives has been discussed in philosophy and jurisprudence since antiquity, without a definitive answer being found. There exists today a vast range of literature¹⁸, nearly overwhelming in its sheer volume, even for experts, which could be a fruitful resource for developing solutions to present problems. It is certainly not the case that the problem should be seen as resolved. In particular, the decision of the Federal Constitutional Court cannot be viewed as the final answer in the debate on conflicts of the type "balancing lives against lives" in emergency situations. Particularly in the criminal law, many questions are still unresolved¹⁹. It would be equally incor-

-
- 17 The shooting down of a passenger plane would be a violation of Article 1 (1) GG as well as the prohibition of killing which derives from it. "This does not change the fact that this approach is intended to protect and preserve the lives of other people." On this subject (written before the BVerfG decision), see Lindner, *Die Öffentliche Verwaltung* 2006, p. 577 *et seq.*
- 18 In addition to the texts cited in footnotes 10 – 17, cf. Archangelskij, *Das Problem des Lebensnotstandes am Beispiel des Abschusses eines von Terroristen entführten Flugzeuges*, 2005; Bott, *In dubio pro Straffreiheit?*, 2011; Coninx, *Das Solidaritätsprinzip im Lebensnotstand*, 2012; Fritze, *Die Tötung Unschuldiger*, 2004; Ladiges, *Die Bekämpfung nicht-staatlicher Angreifer im Luftraum*, 2008; Merkel, *Juristenzeitung* 2007, 373 *et seq.*; Mitsch, *Golddammers Archiv für Strafrecht* 2006, 11 *et seq.*; Pawlik, *Juristenzeitung* 2004, 1045 *et seq.*; Roxin, *Zeitschrift für internationale Strafrechtsdogmatik*, 552 *et seq.*; Sinn, *Neue Zeitschrift für Strafrecht* 2004, 585 *et seq.*; Stübinger, *Notwehr-Folter und Notstands-Tötung*, 2015; Wilenmann, *Zeitschrift für die gesamte Strafrechtswissenschaft* 127 (2015), p. 888 *et seq.*; Zimmermann, *Rettungstötungen*, 2008; Zoglauer, *Tödliche Konflikte. Moralisches Handeln zwischen Leben und Tod*, 2007. Even in older writings, cases of "life-balanced against-life" decisions, in the context of emergencies, were only rarely regarded as justified, cf. Klefisch, *Monatsschrift für Deutsches Recht* 1950, 258. For more on the legal history of how this problem has been dealt with in German legal doctrine, cf. Wilenmann, *Zeitschrift für die gesamte Strafrechtswissenschaft* 127 (2015), p. 888 (893 *et seq.*). For Anglo-American writings on the "trolley problem" compare the citations above in Fn. 14.
- 19 Quite rightly Schneider wrote in *Münchener Kommentar zum StGB*, 2017, preliminary remarks on § 211 *et seq.*, paragraph 29: "The criminal law principles of the prohibition on the quantification and qualification of human life, as well as the in-

rect to give up on the problem as practically irrelevant or unsolvable, and to push it to the side with a shrug of the shoulders.

II. Ethical and legal guidelines as well as a proposed solution

1. “Setting off” human lives vs. a humane orientation in the law

If it was allowed to “set off” human lives against one another, one could argue that it would be permissible to kill an innocent person, if only in that way could the lives of several other (i.e. more) people be saved²⁰. This would mean, for example, that a vehicle approaching an accident situation at high speed and threatening to kill two severely injured people lying on the road, could or even should swerve to avoid running over those injured persons, even if another person was killed by the evasive maneuver (e.g. someone walking along the side of the road). The justification for programming a system to do that, would, however, contradict a principle inherent in humanely oriented legal systems, namely that human beings and their dignity constitute the “highest value”²¹. This excludes the possibility of “setting off” human lives against other human lives, according to the overwhelming view in German legal science and court jurisprudence²².

commensurability of the value of life, are among the frequently highlighted but rarely verified basic convictions of criminal law practice and criminal legal science.”

20 This position is often attributed to utilitarianism, but as a rule no particular representative of this school of thought is named. A strict “set-off solution” would probably be justified from the viewpoint of a less reflected act utilitarianism, but in contrast would not be justified from the perspective of rule utilitarianism. Utilitarian arguments are usually much more sophisticated than is characterized in discussions by German speakers. For a provocative treatment, cf. Peter Singer, *Neue Zürcher Zeitung* 24.5.2015, who even wants to “set off” the lives of pigs against the lives of human beings (<http://www.nzzas/nzzas-am-sonntag/philosoph-peter-singer-ein-embryo-hat-kein-recht-auf-leben-1.18547574>). This suggestion violates two taboos: setting off lives against each other and weighing human lives against animal lives.

21 The concept of “highest value”, like the word “innocent” (footnote 16), needs clarification. It is used here to describe the notion that the legal order is intended to serve the individual whose dignity cannot yield to any other exigency such as “people”, “class” or “will of the God”.

22 OGHSt 1, 321 (334); BGH, *Neue Juristische Wochenzeitschrift* 1953, 513 (514); BGHSt 35, 347 (350); Kühl in Lackner & Kühl, *Strafgesetzbuch*, 2014, § 34 StGB,

On the other hand, it would be very difficult, both morally and legally, in emergency situations, in which the killing of innocent people cannot be avoided, not to try to injure as few innocent people as possible. Therefore, a quantification of victims hardly seems avoidable. In any case, it would not be morally convincing to assert that morally there is no difference between the killing of one innocent person, or respectively, the killing of several or even many innocent people. Suppose a misanthropic programmer wrote a collision algorithm so that his vehicles always killed the largest possible number of people in “set off” situations. Such an algorithm would hardly be regarded as morally acceptable, because we intuitively demand that the number of innocent people killed be kept as low as possible. Perhaps even more counterintuitive would be a system which, appearing to follow the principles that lives cannot be set off against each other and that the destruction of one innocent life is just as “bad” as the destruction of very many innocent lives, was programmed with an algorithm so that in potential accident situations the vehicle killed the lowest possible number of people when it was south of the Main River (i.e. within Bavaria), but in contrast killed the largest possible number of people when it was north of the Main River (outside Bavaria). Such a “Bavaria friendly” collision algorithm should get even the most stubborn set off skeptics to start ruminating²³.

On the basis of legal humanism²⁴, it appears necessary to keep the number of victims as low as possible in cases of the unavoidable killing of innocent people. Therefore if an automatic collision system is faced with the choice between killing one or several innocent people, in a situation where an accident is unavoidable, the avoidance system should choose the solution in which only one, and not several, persons are hit by the car. It is likely that this result will correspond to the moral intuition of most people,

paragraph 7; Perron in Schönke & Schröder, *Strafgesetzbuch*, 2014, § 34 StGB, paragraph 23; Roxin, *Strafrecht AT I*, § 16 paragraph 29; Welzel, *Zeitschrift für die gesamte Strafrechtswissenschaft* 63 (1951), 47 (52); cf. also Ladiges, *Juristische Schulung* 2011, p. 879 (882 *et seq.*), who discusses the problem in the context of “legal justifications” for killing a human being.

- 23 If one analyzes the reasons for our intuitive rejection of such an algorithm, one major factor seems to be the fact that using geographic location as the distinguishing criterion violates our notions of human equality, that is, it is an unacceptable or invalid criterion.
- 24 Hilgendorf, “Humanismus und Recht – Humanistisches Recht? Eine erste Orientierung“ in Groschopp (ed.), *Humanismus und Humanisierung*, 2014, pp. 36 – 56.

that is, prevailing social ethics. It still needs to be investigated whether this principle can stand as it is, or whether it needs further refinement.

2. A proposed solution: Degrees of wrong

According to the view developed here, the killing of innocent people should always be unlawful, even in emergency situations. Let's go back to the hypothetical situation we discussed earlier. Please recall that a car with an automatic collision avoidance assistant was rapidly approaching an accident site, in which three persons had been thrown out of a car and lay seriously injured on the ground. One person was able to drag himself to the side of the road and was leaning on a sign post. The approaching vehicle faced the "decision"²⁵ either to stay in its lane and run over the three injured people, whereby it would have been very likely that all three of them would be killed, or to swerve to the right and kill the fourth person standing at the roadside. If that happened, a human driver would not be able to rely on the justification defence contained in § 34 German Criminal Code (*Strafgesetzbuch* – StGB): the emergency situation – the expected outcome of killing the three injured people on the ground – cannot be legally avoided by changing the trajectory of the car so that it then causes the death of the individual standing at the roadside. An assessment of the legal interests of the parties, based on what has up to now been prevailing legal opinion, would determine that the one protected legal interest, i.e. the lives of the three injured persons, did not significantly outweigh the other legal interest, which would be prejudiced by the maneuver, i.e. the right to life of the individual who would be killed²⁶: Even three lives do not "count" for more than one life when legal interests are balanced, since each individual life in and of itself represents the highest possible maximum value. It should be noted that, strictly speaking, this result was not achieved by prohibiting the balancing of "lives against lives", but rather

25 Once again, the question has arisen whether a hitherto anthropocentric vocabulary can easily be applied to machines ("autonomous actors"). On this subject, cf. Hilgendorf, "Können Roboter schuldhaft handeln? Zur Übertragbarkeit unseres normativen Grundvokabulars auf Maschinen" in Beck (ed.), *Jenseits von Mensch und Maschine*, 2012, pp. 119 – 132.

26 This is the almost unanimous view, cf. Kühl in Lackner & Kühl, *op.cit.*, 2014, § 34 StGB, paragraph 7.

by stressing one very specific consideration: No life counts more than any other life, and even the lives of many persons cannot be classed as more valuable than the life of a single individual. From a humanist perspective, the individual and his rights are the guiding values in our legal order, so that as a matter of principle it is not permissible to oblige the individual to sacrifice his life or other basic rights for the benefit of others²⁷, that is to say, to tolerate being killed or being the victim of serious bodily harm in the furtherance of the interests of others.

The orientation towards the life of the individual as a “non-balanceable highest value” can be explained by the fact that after 1945, the drafters of the German Constitution consciously chose man and his individual dignity as the point of reference and goal of the entire legal system²⁸. This is especially clear from the sentence which was proposed as the first sentence of the Article 1(1) of the draft constitution at the Herrenchiemsee Constitutional Convention (10-23 August, 1948): “The state exists for the people, not the people for the state.”²⁹ The human orientation of law, expressed in this way, is an essential element of the approach to the rule of law established in the German Federal Constitution. The principle of human orientation has been developed into a “humanist imperative” in our law: Na-

27 For an apt treatment, cf. Erb, *Münchener Kommentar zum StGB*, § 34 paragraph 116, 2017, with additional citations, who discusses the “absolute limits of a duty” to sacrifice one’s own life for the benefit of others and rightly wants to apply this principle to “serious health problems”. For arguments in the same direction, cf. Frister, *Strafrecht Allgemeiner Teil*, 2015, chapter 17 paragraph 15; Wilenmann, *Zeitschrift für die gesamte Strafrechtswissenschaft* 217 (2015), p. 888 (909).

28 Life is for this reason unpredictable, because the acceptance of substantial losses of rights in respect of formally protected legal interests cannot be demanded of citizens, except where they are responsible for those losses. For the rules which permit invasions (exceptions) to protected legal interests must be justifiable from the perspective of the individual “(written so concisely by Wilenmann, *Zeitschrift für die gesamte Strafrechtswissenschaft* 127 (2015), p. 888 (909)).

29 Quoted after Dreier, *Grundgesetz Kommentar*, 2013, Art. 1 (1), paragraph 23. Translation by W. Schäubele, “The Herrenchiemsee Constitutional Convention 60 Years On”, a speech held by the Interior Minister on 7.20.2008. It must be pointed out, however, that the GG and its interpretation by the Federal Constitutional Court also includes approaches, which relativize the orientation of the legal order towards the individual, by means of the concept of “community-relatedness” (i.e. public welfare), cf. BVerfGE 4, 7 (15). For more detail on the unpredictability of reference to the “image of mankind” in the GG, see Hilgendorf, “Konzeptionen des ‘Menschenbilds’ und das Recht” in Joerden *et al.* (ed.), *Menschenwürde und Medizin. Ein interdisziplinäres Handbuch*, 2013, p. 195 – 216 (203 *et seq.*).

tional law has to ensure that despite all societal, economic, scientific and technological developments, the individual human being remains the centre and starting point of the entire legal system.

The significance of this humanistic postulate becomes clearer if one compares it with conceptions of the state present in other forms of government: In a *theocracy*, man, with his needs and wants, is not at the centre of the law, but rather it is the will of the deity, which lays claim to obedience to its precepts even when they are associated with the greatest possible human suffering. Man is no more than a “slave of God”³⁰. Another decidedly non-humanistic form of government is the totalitarian dictatorship, following the Stalinist or National Socialist model, in which the rights of the individual are completely suppressed. As a rule, totalitarian governments try to legitimize themselves by citing overarching and pressing national needs (reasons of state) or the needs of the collective (“It is good to die for the Fatherland”, “You are nothing, your people are everything”)³¹.

In contrast, according to our current humanistic legal understanding, which from the rise of Humanism in the 16th century up to the Enlightenment of the eighteenth century largely gained acceptance in Europe, the individual, with his dignity and his “innate” human rights, is at the centre of the legal order. In modern times, this position was formulated for the first time in the early 16th century by authors such as Pico della Mirandola³². Of course in intellectual history terms, it can be traced all the way back to the Greeks and Romans of antiquity³³. This central position of the individual would be jeopardized if the life of one person could easily be set off against the lives of others in emergency situations. The prohibition on setting off lives against each other is based on considerations of principle, which do not change, even where large numbers of lives are at stake. Thus an individual human life cannot be balanced against the lives of 100, 1,000 or 100,000 other people; killing one to save the many is still unlaw-

30 On this theme, which is found both in Christianity and in Islam, cf. Hattenhauer, “Die Sklaven Gottes” in Finkenauer (ed.), *Sklaverei und Freilassung im römischen Recht, Symposium für Hans Josef Wieling zum 70. Geburtstag*, 2006, p. 59 – 82.

31 Regarding the latter, cf. Stolleis, *Gemeinwohlformeln im nationalsozialistischen Recht*, 1974.

32 Pico della Mirandola, *De hominis dignitate, Über die Würde des Menschen* (1496), 1990 (Philosophische Bibliothek vol. 427).

33 Cancik, “Freiheit und Menschenwürde im ethischen und politischen Diskurs der Antike” in Cancik (ed. Cancik-Lindemaier), *Europa – Antike – Humanismus. Humanistische Versuche und Vorarbeiten.*, 2011, p. 175 – 189.

ful. However, in the case of people who are confronted with such horrible decisions, their unease will grow as the number of innocent victims rises, and ultimately, they will indeed decide to kill the lesser number of victims. This psychological reaction can be taken into account in the criminal law by a legal exculpation: the action is illegal, but where the danger cannot be avoided in any other way, the actor incurs no criminal liability under the defences of exculpatory emergency (§ 35 StGB) or extra-statutory exculpatory emergency³⁴.

An argument against the basic position represented here could be seen in the fact that it is based both on a particular conception of human nature which was developed in Europe, as well as on specific understanding of human dignity; it is not intuitively obvious and may require more detailed and compelling reasons. This objection, which is often found in the philosophical debate on human dignity³⁵, strikes the issue at its core. The idea of man as a unique creature endowed with dignity, is a product of the European intellectual history which began in ancient Greece: the pointed emphasis on human dignity after World War II, which among other things was expressed in the prohibition against setting off human lives against each other, was a reaction to the unprecedented crimes against humanity committed under National Socialism in Germany (and Stalinism in the Soviet Union). It does not follow from the historically very specific way in which this particular conception of human dignity emerged, that it is necessarily correct or valid. Of course the values expressing themselves in the

34 Necessity not envisaged by the law, for example, was pleaded as a defence in criminal proceedings to charges of murdering mentally ill patients during the 3rd Reich; the heads of asylums claimed that they had let a certain number of innocent patients be killed in order to save a considerably larger number of patients (see the references in Fn. 12). In its decision on the Air Safety Act, the Federal Constitutional Court alluded to a similar solution in the case of a passenger aircraft hijacked by terrorists, which would be shot down by the German military before reaching the intended location at which it would be used as a weapon of mass destruction. This result was dealt with in a literary context in Ferdinand von Schirach's play *Terror* (2015), whose treatment of the legal issues, however, was not entirely convincing. After the broadcast of the filming of the play on 17.10.2016, according to press reports, about 86% of the viewers voted for "acquittal" of the soldier performing the execution. For criticism of Schirach's play from a legal perspective, cf. Schild, *Verwirrende Rechtsbelehrung, Zu F. von Schirachs 'Terror'*, 2016.

35 Hilgendorf, "Menschenrechte/Menschenwürde" in Cancik *et al.* (ed.), *Humanismus: Grundbegriffe*, 2016, p. 275 – 288 (285 *et seq.*).

“prohibition on setting off lives” are an essential aspect of a humanistic understanding of the law, which has been put forward since the Enlightenment with the claim of having universal validity. This conception was taken up in the German Federal Constitution (1949), and, as enshrined in Art. 1 GG, it is a mandatory rule of our constitutional order.

If the killing of innocent people is always unlawful on the basis of a humanist understanding of the law, the question arises as to how the idea of minimizing the number of lives lost, as above, can be justified in cases where life inevitably is at stake. According to our view, in cases where the dilemma of “weighing life against life” arises, when a decision has to be made between the destruction of one life and the destruction of another life, the principle of the lesser evil must always be followed: if innocent people must die, then it should be as few as possible. If nothing else, this follows from the superior position of the individual developed above. Otherwise, we would consider the two surplus lives, so to speak, as a *quantité négligeable*. The killing of every innocent person remains wrong and cannot be justified. One ought, however, apply the notion of *degrees of wrong*³⁶, which dictates that one should put as few lives as possible at risk, or indeed cause as few deaths as could be possible.

This position can be illustrated by the following hypothetical example: During an airplane crash, the pilot can either steer the plane so that it crashes over a nearly uninhabited area (so that only he himself and all his passengers plus a few people on the ground are killed) or steer it so that the machine crashes over a densely populated area so that not only everyone on board the plane is killed, but also a few hundred or a thousand people on the ground will almost certainly be killed. According to the approach outlined above, the pilot not only has a moral duty but also a legal duty to steer the plane so that the crash takes place over the sparsely populated area. The argument that when human beings are killed every quantification or balancing of lives is impermissible, because there is no normatively relevant difference between the killing of a few persons or many persons, is not convincing because it reduces human life to a *quantité négligeable*. Every human life counts! It remains the case that the killing of innocent human beings is not condoned by the legal system, but rather is classified as a wrong. It follows that the potential victims on the ground

36 Hilgendorf, “Recht und autonome Maschinen – ein Problemaufriß” in Hilgendorf & Hötitzsch (eds.), *Beiträge der 1. Würzburger Tagung zum Technikrecht*, 2015, pp. 11 – 40 (26).

would retain a right of self defence in the moments prior to the crash. On their part, it would be lawful, therefore, if they tried to shoot down the approaching aircraft.

3. *Use of deadly force in especially grave emergency situations involving or not involving risk communities*

Looking at the issue of “risk community”, it is necessary to discuss what role it can or should play that the persons who are at risk, or are threatened, all equally face the same risk. *Gefahrgemeinschaft*, which translates as “risk community”, is a term in use in German law. It is defined as a group of persons facing a certain risk, which may be aware that it is facing that risk. Such a situation would occur, for example, when in heavy rush hour city traffic three children A, B and C suddenly jumped in front of a vehicle in such a way that, without the car being able to swerve, two of the children (A and B) would be hit by the car’s right fender, but the other child (C) would be hit by the left fender. Had the driver had been able swerve the car, he could have steered it to hit either A and B, or C (it was not possible to completely avoid the accident by braking or swerving).

In situations where a risk community exists, as described in the previous paragraph, it must first of all be stressed again³⁷ that neither the killing of A and B nor the killing of C can be justified. Nevertheless, the question still arises as to whether the car should simply drive straight forward, without swerving – then hitting all three children – or swerve to the right, resulting in a collision with A and B, or swerve to the left with the consequence of a collision (only) with C. It seems to me that this hypothetical problem certainly provides support for making decisions based on degrees of wrong³⁸: it is ethically and legally necessary in order to minimize the injuries caused to swerve the car hit C rather than A and B, if the same probabilities of injury and severity of expected injury are present (death, severe bodily harm). Simply invoking “destiny” or the “will of God” as a justification for driving straight ahead and killing all three children seems just as unconvincing as making a decision based on spurious criteria such

37 See above, p. 65 et seq.

38 See above, Fn. 36.

as age, gender or skin color. It would be equally absurd to swerve so that A and B would be struck and killed, for example using the argument that human lives cannot be quantified or balanced against each other, and therefore it does not (from a normative perspective) make any difference whether one, two or three children are killed.

Not convincing (but possibly sustainable) would be calls that a random number generator be used to make the decision on behalf of the programmer. But suppose the random decision was that all three children be killed, although two of them could have been saved – would such a decision be compatible with the fundamental values of our legal system? And who could convincingly make the case to the parents that it was the right decision? Moreover, failure to devise an algorithm based on a hierarchy of outcomes, contains the implicit decision, for which we are responsible, that the result should be left to chance. Finally, it would also be conceivable to open up the possibility for each respective driver to determine in advance how his vehicle will behave in collision scenarios, such as those discussed here, within a range of predetermined possible outcomes. It is obvious, however, that the ethical and legal problems discussed here have not been resolved, but only put off for later.

It should be borne in mind that according to the linguistic usage proposed here, a risk community not only exists when the legal interests concerned have already been massively and specifically put at risk. It is sufficient if the affected legal interests were in principle put at the same risk. One might therefore call it a “symmetrical risk community”.

It still needs to be settled how cases will be dealt with where the potential victims of the collision, at the point in time when the computer system makes its decision, do not face the same risks (i.e. there is no risk community). This would be the case, for example, where the car is approaching a group of three seriously injured people (lying on the street) while a single individual is standing at the side of the road, who would certainly be struck and killed if the vehicle swerved to avoid the persons lying on the road³⁹.

According to the approach presented here, there is no justification for the killing of innocent persons, no matter what decision the system makes. Whatever transpires will be wrong. If one assumes that the risk of being killed is the same both for the three persons lying on the road and for the

39 See above, p. 57.

single individual standing at the side of the road, then one could once again argue that the principle, by which the greatest possible number of innocent persons should be saved, ought to be adhered to so that the vehicle should swerve to avoid the three severely injured people in the road, thereby killing the individual at the side of the road. But that would ignore the fact that before the computer made its decision, the chances of survival were not equally distributed. The vehicle was driving towards the three severely injured people on the road. It only threatened to kill them. If the vehicle is caused to swerve, then the chances of survival are being changed (redistributed). In accordance with social morality (which is not quite clear in this case⁴⁰), there is much evidence here that such a redistribution of the chances of survival should be regarded as incompatible with the humanistic principle of an orientation to the human being as a maximum value of our legal order. The final result of this case is significantly different from that of the risk community, in which all legal interests concerned faced the same risk before the decision was made by the computer⁴¹. In the instant scenario the algorithm should therefore be designed so that the vehicle does not swerve⁴².

III. The quantification of human life in current applicable law

The position developed here contradicts the often somewhat thoughtlessly made assertion in Germany that human life cannot be quantified or at least

40 In the context of the trolley problem, the variant discussed here would probably correspond to the “fat man problem”, cf. Edmonds, *Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us About Right and Wrong*, 2014, p. 35 *et seq.*

41 In the airplane scenario as well, one could well assume “normatively equal” risks, if the aircraft was still far away from possible crash targets.

42 The present paper does not deal with problems of evidence and computer errors. In order to avoid evidential difficulties, black boxes should be installed in all vehicles with high degrees of automation. Computer errors also present an interesting problem: What are the legal consequences, when a computer incorrectly records or misinterprets data? Instead of a disparity between “imagination and reality”, the discrepancy here is between “internal representation and reality”. In the present state of AI research, however, there seems to be good reason to ignore factual errors and mistakes of law by machines, since categories such as “wrong” and “guilt” can hardly be applied sensibly to machines. Cf. sources referred to in Fn. 25.

must not be quantified. In the former assertion, the proposition is obviously wrong: the fact is that human life can be quantified. Anyone can see that this is so by counting the number of living people in a group of human beings, that is, determining their quantity. What is meant by the proposition is not that it is factually impossible to quantify human beings, but rather that it should be forbidden without exception: human lives *should* not, and *must* not be quantified in contexts in which the killing of human beings is an issue under discussion.

There also exist *de lege lata* areas, in which a quantification of human life is permitted, or indeed even required. One example would be sentencing (§ 46 StGB) or aggravating factors at sentencing as under § 306 b (1) StGB. It should be obvious that the sentence imposed by the court will be different where the offender has killed one or many people. This is an issue which must be addressed during the sentencing procedure. We may even go one step further and say that the more people who have been killed, the longer will be the sentence imposed on the offender by the court. A second, and less clear area in which quantifying considerations may play a role is conflicting duties, for example in cases where only one or more persons can be rescued at the cost of others being sacrificed. Let's look at a hypothetical case: a lifeguard has to choose between rescuing child A or rescuing the group of children B, C, and D. Should he not be required to save the group of children rather than the individual child? This question has not yet been conclusively resolved in German legal doctrine.⁴³

A further area in which the quantification of human life is permissible is within the context of the application of the principle of proportionality. Let's look at a hypothetical example: a necessary police operation can be carried out in two equally effective ways, a and b. In operation a, the life of only one person who is not involved would be endangered, but in operation b the number of innocent people put at risk is five. It seems obvious that a quantification of the human lives put at risk by the respective operations, is not only permissible but must be conducted. Should not the same also be true when the innocent people are not only put at risk, but where it is certain or nearly certain that some of them will be killed? A quantification of human life also appears to be necessary, not where innocent people are involved, but where perpetrators will be affected or killed. If the police

43 Cf. Merkel, *Juristenzeitung* 2007, 373 (380).

have various different ways of preventing a terrorist attack (in which innocent people would be killed), they cannot simply kill all the terrorists (by dropping a bomb on them), but must choose the safest measure, i.e. the one which harms or kills the lowest number of victims possible, even if the victims are terrorist attackers. Such an operation would therefore only be legal if conducted in the way least likely to cause lives to be lost (possibly having quantified and compared the possible numbers of victims likely to be caused in the various scenarios under consideration).

Finally, the quantification of human lives during wartime also needs to be touched upon. Is it permissible for a commander of troops to send those troops to their certain deaths in order to save a larger number of human lives? In Hollywood films volunteers are solicited who know their chances of survival are negligible. In the legal literature one certainly does find authors who maintain that an order sending soldiers to their deaths is a lawful order in German law under § 11 (1) of the Soldiers Act (*Gesetz über die Rechtsstellung von Soldaten* – SoldatenG)⁴⁴. Finally, the quantification problem is also discussed in the context of the distribution of scarce resources in the health care sector (i.e. medical triage decisions).⁴⁵

IV. Special problems

We still need to discuss whether other factors must be taken into account, when weighing “life against life”, in addition to the factor quantity (in symmetrical risk community cases). According to the normative requirements of the German Federal Constitution (*Grundgesetz* – GG), factors such as age, gender, ethnicity, health, etc., are from the outset not considered relevant factors.

44 On this discussion, see for example Eser, “Töten im Krieg: Rückfragen an das Staats- und Völkerrecht” in Appel *et al.*, (eds.) *Öffentliches Recht im offenen Staat. Festschrift für Rainer Wahl zum 70. Geburtstag*, 2011, p. 665 – 687 (675 *et seq.*); Leisner, *Das Lebensrecht*, 1976, p. 38, who even considers obvious “suicide missions” to be legitimate, provided that many people could be saved.

45 Giesen, *Juristenzeitung* 1990, 929 (941 *et seq.*).

1. The probability of being injured

There is very much to be said for taking into account the factor “probability of being injured”. Who should the computer system decide to run over, therefore, if the two seriously injured people on the ground are very unlikely to be killed by the approaching car, but the pedestrian at the roadside with the greatest certainty would be killed, if the vehicle swerved and hit him? Our ethical intuition speaks in favour of taking into account the probability of injury when weighing the interests of the parties involved. The law also requires that it be considered, when § 34 StGB focuses on the “...the degree of ... danger facing them ...” It seems, however, that we have reached the limit of what can be meaningfully asserted given the present state of our knowledge. It seems very unlikely in the foreseeable future, that it will be possible to accurately quantify the probability of injury in real (i.e. not hypothetical) accident situations. At best it will be possible to make qualitative or comparative statements, i.e. statements such as “almost certain”, “very likely”, “very unlikely” or statements such as “event A is more likely than event B”.

In the absence of “hard” probabilities, it is not possible to formulate clear, unambiguous rules for dealing with relevant conflicting interest scenarios. Given the choice of either (a) the certain (or near certain) killing a person or (b) placing one or two persons into situations where the risk of death is low, a decision in favor of choice (b) would probably be in accordance with the moral intuition of most people. In the final analysis, therefore, the majority of arguments are in favour of including the likelihood of injury into the weighing of interests in the dilemma situations in question, also and especially when lives are being weighed against lives.

2. Self-protection measures

How self-protection measures will be fed into the equation remains to be clarified: Suppose a car is in an emergency situation in which it is impossible to avoid a collision. The victim will be one of two cyclists. Should there be a preference for the car to hit the cyclist wearing a crash hel-

met?⁴⁶ This would take into account the fact that this person is better protected against injuries in the event of a collision with the car. On the other hand, it would mean that those road users, who try to take appropriate measures to protect themselves, would be treated less well than those who recklessly refuse to use protective devices like helmets. Should the vehicle be steered to hit the cyclist without a helmet, because he has refused to take appropriate safety precautions? This highlights a fact that has often received little attention in the debate so far: From the perspective of those persons who are directly affected, a system designed to avoid collisions can behave like an attack system⁴⁷.

The special problem results from the fact that the necessity of minimizing risks and injuries, if possible, i.e. the principle of the lesser evil, conflicts with aspects of prevention. Strictly speaking, potential damage should be minimized and therefore the cyclist with the crash helmet should be hit by the car. This, however, would provide an incentive not to wear a crash helmet, i.e. refrain from taking protective measures in road traffic, which hardly seems acceptable from the perspective of injury prevention. Based on the humanistic understanding of law expressed here, we would have to insist that operating a motor vehicle in such a way that it would kill or almost certainly kill victims who failed to use protective devices is certainly not permissible. Educating road users to use protective devices such as helmets should certainly not be an issue here. In all other cases, the goal should be to maneuver the vehicle in such a way as to make a collision less likely, disregarding the extent to which the victim is wearing protective gear. This approach, however, is only one of several legal policy options which may appear to be acceptable.

3. *Actions and omissions*

One issue, which could be very important from the point of view of the criminal law, is the distinction between actions and omissions. Not only do offences of omission, in contrast to offences committed through actions,

46 In principle, the same question arises with regard to safely designed and less safely designed cars and their passengers (who, for example, may or may not be wearing their seat belts). Cf. also below section 3.

47 Lin, "Why Ethics Matters for Autonomous Cars" in Maurer *et al.* (eds.): *Autonomes Fahren*, 2015, pp. 69 – 85 (72 *et seq.*).

require that additional criteria be fulfilled for criminal liability to be incurred, but according to prevailing German legal opinion it may be easier in the case of homicide offences committed by omission to avoid liability through the use of legal justifications than is the case for offences committed through actions. In particular, according to prevailing opinion, in cases where there are conflicting duties, the legal interest which the actor subject to the conflicting duties decides to protect, does not need to be “significantly more important” than the legal interest he chooses not to protect. Rather, for the legal justification to apply it is sufficient that two equally important legal interests are at risk, and the actor can only protect one of those interests, so he chooses from the two, and protects that legal interest⁴⁸. Let’s look at a hypothetical example: As in our previous case, a vehicle is travelling at high speed toward a situation in which three seriously injured persons A, B and C are lying in the road in its path. It does not swerve to avoid running over the injured people on the road, because if it did so it would collide with and kill D, who is standing at the side of the road. One could argue that A, B, and C’s deaths were not caused by an action of the driver, but rather through the omission of the driver, namely his failure to swerve to avoid running them over⁴⁹. This would be a way to interpret the facts in order to “create” a potential offence of omission, in which under certain circumstances a legal justification through conflicting duties might arise.

This argument, however, is not convincing for several reasons: A vehicle that simply travels straight forward and collides with a person, without the driver steering to alter the trajectory of the vehicle, harms the victim (here the person struck by the car) through the action of running him over. If it were otherwise, a large proportion of road traffic offences would be offences of omission rather than offences where the perpetrator performed a positive act. The fact that the car could have swerved to the right or left does not change this conclusion⁵⁰. In addition, it should be borne in mind that cases where a vehicle simply moves “straight forward” are probably more the exception than the rule. It is equally conceivable that the vehicle

48 Cf. Neumann in Kindhäuser *et al.*, (eds.) *Strafgesetzbuch* (Nomos-Kommentar), 2017, § 34 StGB, paragraph 124 *et seq.*

49 From a purely logical point of view, of course, this argument would be acceptable.

50 Of course a marksman can intentionally miss his target, by shooting to the right or to the left. That does not mean, however, that intentionally missing the target should be regarded as an omission (not hitting the target).

can only turn to the left (injuring person A there) or to the right (injuring person B there)⁵¹.

Furthermore, it is true that in German criminal proceedings a collision of duties can be used as a legal justification by a (human) actor who might otherwise have incurred criminal liability. However, it is certainly not clear that this legal concept can also simply be applied to the actions or omissions of machines controlled by algorithms! More likely would be an assessment of the situation at issue by a court from the perspective of the victim and his fundamental rights. From the perspective of the targeted (innocent) human collision victim, the “behaviour” of the vehicle would be interpreted as an unlawful infringement of his fundamental rights to life and physical integrity (Article 2 (2)(1) GG). Therefore, in the conflict scenarios sketched out above, neither of the alternatives achieves a satisfactory result.

V. The liability of manufacturers of collision avoidance systems

The questions discussed so far have concerned the assessment of concrete emergency situations. This has to be distinguished from two questions: (1) whether manufacturers can be held liable for collision avoidance systems, if property damage or personal injury occur; (2) whether automatic collision avoidance systems, together with their respective programs, should or should not from the outset be licensed for use on public roads because of the risk of unlawful fatal accidents⁵².

1. Exclusion of liability using the concept of “accepted risk”

According to the view expressed here, such systems are permissible and their introduction is necessary and desirable. In order to avoid the risk of civil liability but in particular to avoid the risk of criminal liability, it is

51 For example, at a fork in the road, etc.

52 Rejection of their use on public roads, however, would again have to be justified both morally and legally – a pathway back to the state of innocence, before collision avoidance algorithms were technologically feasible, does not appear to be possible. This demonstrates how our technological capabilities not only extend our actual possibilities, but also create responsibilities.

necessary to do everything possible and reasonable during the programming and installation of computer systems, working at the state of the art of the technology, in order to avoid causing damage. Subsequently, the systems must also be monitored, serviced at regular intervals and, if possible, updated⁵³. These requirements follow from the doctrine of accepted risk, by which duty of care requirements are restricted in the case of technologies which are deemed to be fundamentally positive⁵⁴. Just as airbags and seatbelts may be (or even must be⁵⁵) installed and used in motor vehicles although in some cases they can result in personal injury or even death, the installation of automatic collision avoidance systems is not considered to be a breach of duty of care requirements (i.e. negligence) as long as all reasonable technological solutions have been implemented in order to minimize potential injuries. This assertion requires a somewhat more detailed explanation:

Society of the present is marked by the development and the constant introduction of new technologies that are accompanied by new risks. This can be seen in new medicines and in new forms of medical treatment as well as in energy production, food production and road transport⁵⁶. The allocation of liability risks has developed into a core problem for the law of present: “In modern ‘risk society’, interest in the distribution of material goods is being pushed further and further into the background by the more existential concern of how potential risks, which up to now could never have been imagined in their dimensions and ubiquity, should be allocated⁵⁷.”

Not every infringement of a legal interest is a criminal offence. The conscious decision that a risk is an accepted risk may appear sensible if

53 Such improvements are likely to be made in the future, largely by installing improved software either in repair workshops or by radio.

54 Kindhäuser, *Strafrecht Allgemeiner Teil*, 2017, § 33 paragraph 26; Kindhäuser, *Strafgesetzbuch. Lehr- und Praxiskommentar*, 2017, § 15 paragraph 58.

55 On the legal duty to wear a seat belt, cf. § 21 a German Highway Code (*Strassenverkehrs-Ordnung* – StVO).

56 For a more detailed treatment from a sociological perspective, cf. Hoyer, *Zeitschrift für die gesamte Strafrechtswissenschaft* 121 (2009), p. 860 *et seq.*

57 Duttge, *Zur Bestimmtheit des Handlungsunwerts bei Fahrlässigkeitsdelikten*, 2001, p. 489. On the concept “risk society” and its reception in the law, cf. Hilgen-dorf, *Strafrechtliche Produzentenhaftung in der “Risikogesellschaft”*, 1993; Reus, *Das Recht in der Risikogesellschaft, Der Beitrag des Strafrechts zum Schutz vor modernen Produktgefahren*, 2010, both with detailed citations.

the associated positive consequences clearly outweigh the negative consequences. The notion of “accepted risk” suggested here may be found scattered in many disparate areas of the law. Karl Binding, to whom we owe thanks for the first more detailed analysis of this concept, referred 100 years ago to this as “isolated traces of a great legal idea”⁵⁸. In particular, accepted risk was recognized very early in connection with trade and technology. There is a motto associated with the Hansa, a very successful confederation of German market towns and guilds, which controlled the Baltic sea trade in the late middle ages: “*Navigare necesse est, vivere non necesse*”⁵⁹, which can be roughly translated as: “That we go to sea is necessary, that we all survive is not”.

Binding states the idea in a more general way as follows: “The more indispensable an action is in a legal sense, the greater the risk that it can be done without legal repercussions⁶⁰.” The “indispensability” of the action can result from its significance for an actor or his relatives, but also because of the meaning it has “for certain sections of society, or for the legal order and the state⁶¹.” The creation of risks is only permissible, however, as far as is necessary⁶². Jakobs has quite rightly pointed out that the permissibility of a risk is often not simply confirmed through a cost-benefit analysis; rather, besides “accepted risk by risk assessment”, there also exists accepted risk by virtue of “historical legitimation”⁶³. This leads us to the issue of “social adequacy” as a basis for accepted risk⁶⁴.

It follows from what has been said, as Ulrich Weber pointed out, “the legal order takes certain risks, even risks to life and limb are tolerated with eyes wide open. This has been done, for example, as legal approval was

58 Binding, *Die Normen und ihre Übertretung, Eine Untersuchung über die rechtmäßige Handlung und die Arten des Delikts*, vol. 4: Die Fahrlässigkeit (negligence), 1919, p. 436. For a thorough discussion of the negligence problem today, cf. Duttge, *op.cit.* 2001 (Fn. 57).

59 Binding, *op.cit.* 1919 (Fn. 58), p. 437, who refers to Rümelin, *Schadensersatz ohne Verschulden*, 1910, p. 26.

60 Binding, *op.cit.* 1919 (Fn. 58), p. 440.

61 *Ibid.* (Fn. 58), p. 440 *et seq.*

62 *Ibid.*, (Fn. 58), p. 442.

63 Jakobs, *Strafrecht Allgemeiner Teil*, 1991, 7/36 with reference to BGHZ 24, 21 (26 *et seq.*).

64 An example of the historically and culturally based differential treatment of social risks, which can be examined from a legal perspective using the concept of social adequacy, is the lawfulness of selling “hard” alcoholic beverages, on the one hand, and the ban on cannabis, on the other.

given to motorized road transport and to the operation of hazardous installations, but only on the condition that the parties observed strict safety precautions. If these rules are complied with... a negligence claim cannot be raised even if a socially damaging consequence is the result ... for example, a road user is injured or even killed.⁶⁵

These considerations can be directly applied to the issues raised here: The installation and operation of automatic collision avoidance systems may not in principle be regarded as negligent, even if it is clear that such systems (statistically with near certainty⁶⁶) will also result in harm to human beings under very unfavorable conditions. Certainly, the prerequisite for this is that the systems are designed in such a way that the extent of possible damage is kept as low as possible. The safety rules to be observed thereby can be explicitly stipulated in technical rules, but they can also result from the analysis and assessment of individual cases⁶⁷.

This does not mean that the injury or even killing of an innocent person could be justified by the use of an automatic collision avoidance system (i.e. be lawful). This would be contradicted by Art. 2 (2)(1) GG which places human life under special protection. From the fact that reliance and use of automatic collision avoidance systems is not adjudged to be negligent, it certainly does not follow that if such a system were to kill a human being in an extreme case – one might even say a “misadventure” – this killing would be lawful. The person concerned does not have to tolerate his life being put at risk, which ought to go without saying, but can try to avoid the danger or defend himself.⁶⁸

65 Baumann, Weber & Mitsch, *Strafrecht Allgemeiner Teil*, 11th ed., 2003, § 22 paragraph 14. On this subject, see also Duttge, *op.cit.* 2001 (Fn. 57), p. 104 *et seq.*

66 Nevertheless, it would not be convincing to assume that manufacturers or programmers intend to cause injuries, since in spite of the fact that such damage is statistically almost certain to take place, it cannot be known in advance, when, at which place, and caused by whom, such accidents will occur.

67 In general, compliance with technological rules is not the same as fulfilling the duty of care required in road transport, since the duty of care required in individual cases may go beyond what is necessary to comply with relevant technological rules (which may not be appropriate or no longer appropriate). Compliance with technological rules is, in any event, an indication that the necessary duty of care has been fulfilled. For additional detail, cf. Duttge in *Münchener Kommentar zum StGB*, 2017, § 15 StGB, paragraph 138, which includes further citations.

68 It should be noted that reliance on self-defence under § 32 StGB should be ruled out in the absence of a physical attack. There remains, however, recourse to the defence of necessity under § 34 StGB.

The installation and use of powerful motors, the use of a metal body panels, and the authorization of motorized road transport as such⁶⁹, create risks that can in individual cases result in the deaths of human beings. Nevertheless, the companies that sell powerful car motors and steel car body panels, are no more negligent than the state that permits motorized road transport because the benefits that accrue from these actions more than offset the damage they cause. It is obvious that the killing of human beings through road transport is per se not justified. It would be erroneous to derive an obligation to tolerate individual cases in which risk threatens to be realized (i.e. injury incurred) from the social acceptance of the creation of risk.

2. Counterarguments

Engländer has expressed opposition to the limitation of liability by means of the concept of “accepted risk” in the context of automatic collision avoidance systems in road transport⁷⁰. His arguments, however, do not stand up to critical analysis.

According to Engländer, the two characteristic features of accepted risk are, firstly, “as an exception, the general usefulness of an activity which may cause harm”, and, secondly, a “lack of power to avoid causing harm on the part of an actor, for which he himself bears no responsibility, meaning the inability to prevent the result in individual cases (to the extent that the actor is not prepared to completely renounce engaging in the respective activity)”⁷¹. Engländer regards accepted risk as a ground for excluding objective attribution, and not, as I have suggested,⁷² as a means of limiting negligence claims, which accords with prevailing opinion. Perhaps, this is one reason why Engländer confounds “accepted risk” with respect to the responsibility of vehicle users, with the responsibilities of manufac-

69 Cf. the previously cited statement of Ulrich Weber in Baumann, Weber & Mitsch, *op.cit.* 2003, § 22 paragraph 14. (see above, Fn. 65).

70 Engländer, *Zeitschrift für Internationale Strafrechtsdogmatik* 2016, 608 (612).

71 *Ibid.*

72 Expert testimony before the Committee on Economic Affairs and the Media, Infrastructure, Construction and Transport, Energy and Technology of the Bavarian State Parliament, 17th electoral period, 38th meeting, 19.10.2015, p. 50 (available at <https://www.bayern.landtag.de>).

turers for the programming and marketing of their collision avoidance assistants.

Engländer deals firstly with the possibility of vehicle users incurring criminal liability⁷³.

As a possible *actus reus* which could incur criminal liability, he correctly points to “putting the appropriately programmed vehicle into motion”⁷⁴. As a possible element of an offence, here the criminal result, he refers to the “death, bodily harm, damage to property of another road user”⁷⁵. Amazingly, however, at this point Engländer does not discuss issues of intent or negligence, but focuses solely on the concept of accepted risk. If accepted risk is used as an instrument for limiting negligence claims, this question does not arise because the user of a vehicle with a properly functioning collision avoidance system is not negligent when a correspondingly programmed vehicle is used. A user cannot predict the occurrence of a concrete accident situation in which the collision avoidance system would intervene. Thus it can be seen that a key prerequisite to substantiate a negligence claim is missing, so that the question of restricting the duty of care required by law by means of the concept of accepted risk does not arise.

Rather confusingly, Engländer then explains that the vehicle user (!) possesses the “power to avoid causing the harm” he postulated, since the occurrence of the factual result could “very simply” have been avoided by programming the collision avoidance system differently. That is true, but only as regards the manufacturer or programmer, and certainly not the vehicle user. The vehicle user could only avoid the intervention of the collision avoidance system by not starting up the car, which means that he has to renounce the activity for which he is accused in the first place.

The change of perspective, which is not explained at all, makes it difficult to understand Engländer's argumentation, especially since he later affirms the concept of accepted risk, albeit in favour of manufacturers, under certain circumstances⁷⁶! In this case, the issue of “power to avoid causing harm” by reprogramming the system (which only manufacturers or their programmers are able to do) is apparently are no longer important.

A manufacturer who programs his automatic collision avoidance system according to the rules developed in the first part of this paper, may, in

73 Engländer, *op.cit.* 2016, 608 (611 *et seq.*).

74 *Ibid.*, p. 611.

75 *Ibid.*

76 *Ibid.*, p. 617.

our view, rely on the concept of accepted risk. From a legal point of view, the development and installation of such systems into motor vehicles do not represent a misdeed, even if a (subsequent) death caused by such a system constitutes a wrong⁷⁷. The same (approach) applies here as it does to the use of seat belts and automatically opening airbag systems, which similarly in almost all cases protect their users from harm, but do in rare cases cause deaths, without the manufacturers of such systems incurring negligence liability as long as the systems have been designed as safely as possible, given the current state of the art of the technology.

3. *Passenger protection*

The question remains whether the manufacturer, in the present context, in ensuring the safety his vehicles (and especially their occupants) is subject to new ethical or legal restrictions. If one applies the rules developed above, this is not the case: vehicle occupants are (of course) not obliged to acquiesce in being injured or killed.

Collision avoidance assistants, which are installed into motor vehicles, must be programmed in such a way that vehicle occupants are protected under all circumstances. Only the case of the symmetrical risk community is problematic, i.e. where two (or more) persons or groups of persons face the same risk. In our view, the principle of the lesser evil should be applied in such cases.

This can be visualized using two examples. A vehicle is travelling toward a number of persons lying on the road. It can neither swerve nor brake safely to avoid killing those persons. According to the view presented here, the driver is not under an ethical or legal duty to drive in such a way as to destroy his own vehicle (for example, by swerving so as to collide with a concrete pillar or some other self-destructive manoeuvre), even if the number of people saved would exceed the number sacrificed. This is not a case of a symmetric risk community.⁷⁸

The situation is completely different when a car at high speed is approaching a broken down truck carrying explosives, which is blocking the road. The explosion caused by a collision would not only kill the three oc-

77 See Weber citation above (Fn. 65).

78 See above, p. 71.

cupants of the truck but also the driver of a vehicle approaching in the opposite direction. In such cases, where a symmetric risk community exists, according to the view presented here, the vehicle which is about to collide with the truck must in principle attempt to avoid the collision by swerving, even if this would seriously endanger the life of the occupants of that vehicle⁷⁹.

Such cases should, of course, only occur in theory. Manufacturers are responsible for ensuring the lives and the physical integrity of the occupants of their vehicles by installing high performance safety systems. The solution presented here is simply a further incentive for them to continually optimize passenger safety. In addition, it would be as unreasonable for car manufacturers to be under a legal duty to install “self-destruction mechanisms” in their vehicle as it would be to legally require buyers to use such vehicles⁸⁰. Manufacturers are under no legal duty to produce cars which put their “own” passengers into significant danger or even sacrifice the lives of those passengers, but rather motor vehicles should and must be made to be as safe as possible, even if complete safety can never to be achieved.

4. What risks should be considered “accepted” risks?

This leads to the question as to which risks a society regards as “acceptable” and therefore wishes to classify as “accepted” risks. In a democratic state, the answer requires a process of social debate and “negotiation”. Essential variables for answering this question should include objective and verifiable criteria such as levels of possible damage, probability of occurrence, possibilities for prevention, and the issue of whether or not damage is irreversible. In social reality, however, the acceptance of technological risk is shaped by historical contingencies and often hardly reconstructible prejudices and habits. Even risk perception varies considerably from person to person⁸¹. Lawyers are part of society and convey social risk aware-

79 See above, p. 84.

80 On the principle of reasonableness, see Hilgendorf & Valerius, *Strafrecht Allgemeiner Teil*, 2015, § 11 paragraph 90: An action is unreasonable if, as a result, one's own legitimate interests are harmed to a considerable extent.

81 One example from the area of anti-drug policy are the different perceptions of risks associated with alcohol and cannabis. For further detail on the risk debate, cf.

ness into the judicial decision making process (and under certain circumstances into the legislative process). It is the task of a rational analysis of the consequences of technological development⁸², based on empirical research, to accompany and influence the debate on which risks should be considered “accepted risks”.

It could be feasible to introduce a kind of “algorithm seal of technical approval” for automatic collision avoidance systems (and perhaps for other algorithms that have to make particularly risky decisions), i.e. a special approval procedure, which would be required before a system could be put onto the market. The competent authority for the implementation of such a procedure should be a state authority whose work is subject to safeguards generally accepted in states under the rule of law. One could also imagine certification procedures. In this way it should be possible to control and “fence in”, via norms, technological development in the area of algorithms so as to preserve the humanistic imperative of always accepting the fundamentally free individual, with his special dignity as a human being, as the guiding value of our law and jurisprudence on that law⁸³.

Although, according to the view represented here, the installation of automatic collision avoidance systems is not to be adjudged negligent because the risks created by them in a very small number of cases are more than outweighed by the significant utility they provide in the overwhelming majority of cases, there is still one point which needs to be emphasized. This was looked at by Binding⁸⁴: risk creation is only permitted to the extent that it is *necessary* to achieve the intended benefits for society. What we are dealing with here is a criterion that can be empirically tested. Every risk creation, which goes beyond what is strictly necessary, is reprehensible. This means that new technological systems must be designed in such a way as to minimize the risks created by them. We will once again be able to speak of degrees of wrong and of the duty to reduce the wrong

Fischhoff & Kadavy, *Risk, A Very Short Introduction*, 2011; Renn & Zwick, *Risiko- und Technikakzeptanz*, 1997; for a recent contribution, cf. Renn, *Das Risikoparadox. Warum wir uns vor dem Falschen fürchten*, 2014.

82 Grunwald, *Technikfolgenabschätzung. Eine Einführung*, 2nd ed. 2010; cf. also Grunwald, *Technik und Politikberatung. Philosophische Perspektiven*, 2008.

83 Cf. above p. 63 et seq.

84 Cf. above Fn. 58.

to a minimum. In doing that, of course, the general legal principle of reasonableness must be observed⁸⁵.

The results found for collision avoidance systems can be extended to all technological (and non-technological) products: their development and their use are permitted, even if their use in individual cases can lead to unintended damage, provided the risks arising from them can be considered acceptable⁸⁶. It does not matter whether one treats accepted risk as a separate concept within the legal doctrine⁸⁷ of negligence or merely as the *other side* of the duty of care obligation⁸⁸. It is an important principle of modern product liability law or producer liability law⁸⁹.

VI. Closing remarks

The results can be summarized in the following theses:

1. The transfer of human decision making to algorithm-driven technological systems forces us to make processes explicit, which had previously been done without reflection. That means raising them to the level of con-

85 See above Fn. 80. This means among other things, that manufacturers are not obliged to incur expenditures that could jeopardize their economic competitiveness or even their existence. On the other hand, the state is obliged to ensure adequate protection of its citizens, even in the face of technological developments (above Fn. 36), p. 35.

86 Vogel in *Strafgesetzbuch Leipziger Kommentar*, 2012, § 15 StGB, paragraph 279.

87 For example, according to Lenckner & Sternberg-Lieben in Schönke & Schröder, *op.cit.* 2014 Vor §§ 32 *et seq.* StGB, paragraph 107 b.

88 Kindhäuser wrote in “Zum sog. ‘unerlaubten’ Risiko”, in Bloy *et al.* (eds), *Gerechte Strafe und Legitimes Strafrecht: Festschrift für Manfred Maiwald zum 75. Geburtstag*, 2010, p. 397 (404): “Whoever engages in dangerous acts which are accepted, does not violate his duty of care. And vice versa: Anyone violates his duty of care, engages in dangerous acts which are not accepted.” Cf. also Duttge in *Münchener Kommentar zum StGB*, 2017, § 15 StGB, paragraph 139. This language usage corresponds to everyday (German) language. In this context, however, the concept of accepted risk is clearly understood in a much wider sense than in the above text, where its use is restricted to the development and marketing of hazardous products.

89 Cf. also § 3 Product Liability Act (*Produkthaftungsgesetz* – ProdukthaftG), according to which a product is defective “when it does not provide the safety which one is entitled to expect, taking all circumstances into account, in particular its presentation, the use to which it could reasonably be expected that it would be put, the time when it was put into circulation.” Absolute security cannot be expected.

sciousness and subjecting them to analysis. One could almost speak of a *compelling need to explicate* going hand in hand with the “algorithmization” of the world we live and work in. The conflict dealt with here is, in this respect, only one of many decision making situations which must be rethought in ethical and legal terms.

2. In modern road transport, the *principle of the lesser evil* applies in the event of an emergency situation. This is a general principle of our law. It follows, inter alia, that the protection of life and limb must always take precedence over the protection of property.

3. The principle of the lesser evil becomes problematic when serious personal injuries have to be weighed against one another. This is especially true in the context of life threatening situations.

4. The killing of an innocent human being by an automatic collision avoidance system cannot be justified by saving a greater number of lives. Based on legal-ethical considerations, which are ultimately rooted in legal humanism, the individual human being is the maximum value in our legal order. In principle, he must not be compelled to sacrifice his own central (“essential”) legal interests for the benefit of others.

5. Qualitative characteristics such as age, gender or ethnic origin may not play a role in the assessment of emergency situations. In contrast, the probability of violations of legal interests should be taken into account in computer-controlled decision-making processes.

6. When weighing-up road traffic collisions, considerable difficulties are posed by considerations of the safety precautions of respective road users. As a matter of principle, causing death and serious personal injury must if possible be avoided. Personal issues and characteristics may not play a role in accident prevention and avoidance.

7. The thesis that lives cannot be quantified, if taken literally, is incorrect. Such quantifications are, in certain cases, even morally and legally necessary, for example, when applying the proportionality principle.

8. In emergency situations, in which the lives of several people are equally threatened (symmetrical risk community), an assessment of degrees of wrong should be undertaken: The killing of every innocent human being is legally wrong. Nevertheless, the number of innocent victims should be kept as low as possible. This assumes both the quantification of potential victims, as well as respect for the principle of the lesser evil.

9. In contrast, in cases where a symmetrical risk community is not present, i.e. where all persons involved do not from the outset face equal or at least comparable life-threatening risks, those persons who are not at

risk of serious injury or death should not be put at risk. The prohibition on “redistribution of chances of survival” follows from the same humanistic principles discussed previously in the context of balancing lives (cf. 4 above).

10. The development and use of automatic collision avoidance assistants cannot be regarded as negligent, because, although they create certain risks, these risks are more than offset by their practical benefits (so-called accepted risk). Rather they operate like other technological systems both in and outside the context of road transport: If they provide major social benefits, for example, by significantly increasing the safety of road transport, the development and use of such systems are permissible, even if the systems can cause damage in individual cases that cannot per se be justified.

11. Automatic collision avoidance systems, however, must be designed in such a way that the damage they cause is reduced to the absolute minimum. Thus the principle of the lesser evil applies here, too. The issue of reasonableness is important when considering ways to optimize safety. Manufacturers should and must make their vehicles as safe as possible.

12. The concept of accepted risk can be applied in the research, development, and marketing of all sorts of goods, including but not limited to technology based products, in addition to automatic collision avoidance systems. It is a general principle of modern product liability law and producer liability law, which is likely to play a key role in the future clarification of the legal issue of liability for damage caused by automatic systems.

