

# Ethik im autonomen Fahrzeug: Zum menschlichen Verhalten in drohenden Unfallsituationen

Sabrina Eimler, Stefan Geisler, Philipp Mischewski

Hochschule Ruhr West

{sabrina.eimler, stefan.geisler}@hs-ruhrwest.de,  
philipp.mischewski@gmail.com

## Zusammenfassung

Öffentliche Diskussionen zum autonomen Fahren zeigen, dass ein hoher Anspruch daran besteht, dass die Algorithmen in kritischen Fällen Entscheidungen nach ethischen Kriterien fällen. Diese Kriterien für die Vielzahl von denkbaren Verkehrssituationen so zu erfassen, dass sie den Vorstellungen eines größten Teils der Bevölkerung entsprechen, stellt eine große methodische Herausforderung dar. In dieser Arbeit wird untersucht, in wie weit eine überlegte Entscheidung mit dem Verhalten in einem Fahrsimulator übereinstimmt. Dabei wird bei einem großen Teil der Teilnehmer\*innen ein Widerspruch zwischen geäußertem beabsichtigtem Handeln und tatsächlichem Handeln offenbar.

## 1 Einleitung

Eine Erwartung von Verkehrsteilnehmer\*innen bei der Einführung autonomer Fahrzeuge ist die Erhöhung der Verkehrssicherheit durch Reduktion der Anzahl von Unfällen (Lenz & Friedrich, 2015). Es ist unstrittig, dass ein gesellschaftlicher Nutzen vorliegt, wenn die Gesamtzahl von Verletzten und Toten im Vergleich zum manuellen Fahren sinkt, auch wenn keine komplette Unfallfreiheit erreicht wird. Durch Begrenzungen von Sensorik, Algorithmen und Aktuatorik werden Unfälle nie ganz auszuschließen sein (Goodall, 2014), insbesondere da der individuelle Verkehrsraum ein hohes Grad an Komplexität und Unplanbarkeit beinhaltet. So wird immer davon auszugehen sein, dass die Sicht durch Objekte versperrt sein kann oder sich andere Verkehrsteilnehmer\*innen, insbesondere Fußgänger\*innen, nicht an Regeln halten und auch nicht vollständig durch Vernetzung erfasst werden können.

Eine Frage, die weniger technisch als ethisch zu klären ist, ist die des Verhaltens des Fahrzeugs, wenn ein Unfall unvermeidbar ist (Goodall, 2014; Lin, 2015a). Insbesondere untersuchenswert ist die Situation, in der ein Unfall und Schaden zwar nicht mehr verhindert werden kann, aber eine Möglichkeit besteht, zwischen unterschiedlichen potenziellen Geschädigten zu entscheiden, etwa durch das Ausweichen nach links oder rechts. Nach welchen Kriterien soll diese Entscheidung durchgeführt werden? Diese Frage ist im Grundsatz verwandt mit einem klassischen Problem der Ethik, dem Trolley-Problem. U.a. dieses wird im folgenden Abschnitt dargestellt, ebenso Arbeiten, in denen eine Übertragung auf das autonome Fahren stattfindet.

Eine verwandte Frage ist die, ob Menschen in einer konkreten Situation tatsächlich genau so handeln, wie sie es sich im Vorfeld außerhalb des Verkehrsraums von sich selbst wünschen oder auf Basis ihrer ethischen Prinzipien für richtig halten. Dies zu untersuchen ist Gegenstand dieser Arbeit. Hierzu wird zunächst ein Versuchsaufbau konzipiert und eine Studie durchgeführt (Abschnitt 3), sowie die entsprechenden Ergebnisse ausgewertet und interpretiert (Abschnitt 4). Der Artikel schließt mit einem Fazit und einem Ausblick.

## 2 Verwandte Arbeiten

Die Frage, nach welchem Kriterium ein Mensch ausgewählt wird, der zu Schaden kommt, wenn es nicht möglich ist, alle zu retten, ist in der Philosophie und Ethik vielfach diskutiert. Im Bereich der technischen Automatisierung ist das einflussreichste Gedankenexperiment das sog. Trolley-Problem (Thomson, 1976; Thomson, 1985). Die darin betrachtete Problemstellung ist, dass ein Eisenbahnwagen ungebremst auf eine Gruppe von fünf Menschen zufährt, die aufgrund bestimmter Randbedingungen nicht mehr ausweichen können. Wohl aber besteht für eine außenstehende, die Situation beobachtende Person, die Möglichkeit, eine Weiche so umzustellen, dass der Wagen auf ein anderes Gleis fährt, auf dem sich nur eine Person befindet, die ebenfalls nicht ausweichen kann. Die Frage ist nun, ob eine Aktion moralisch erlaubt oder vielleicht sogar geboten ist, die einen Menschen in eigentlich sicherer Position tötet, um fünf andere, die ansonsten sterben würden, zu retten. Varianten existieren bzgl. der äußeren Bedingungen aber auch bzgl. der Charakterisierung der betroffenen Personen, etwa junge vs. alte Menschen. Eine Möglichkeit, selbst derartige Fragestellungen zu kreieren und abzufragen, stellt die Moral Machine (Rahwan et al., 2016) dar.

Übertragen auf autonome Fahrzeuge stellt sich die Frage, ob ein Algorithmus die ursprünglich geplante Fahrtrajektorie zugunsten einer Person oder Personengruppe aber gleichzeitig zu Lasten einer anderen Person oder Personengruppe abändern darf oder sogar soll. Der Vergleich hat aber auch Grenzen, da bei der Algorithmenentwicklung die Möglichkeit besteht ein gesellschaftliches Meinungsbild einzuholen und ohne Zeitdruck und mit Zusatzinformationen zu entscheiden (Nyhol & Smids, 2016). In Deutschland geschah diese Diskussion zu Teilen bereits im Rahmen einer Ethikkommission (BMVI, 2017).

Frison et al. (2016) beschreiben ein Fahrsimulator-Experiment, in dem Proband\*innen verschiedene Situationen gezeigt wurden, in denen sie entweder sich selber oder andere in Gefahr brachten. Je nach Anzahl, Alter und ggf. vorhandene Beziehung zur gefährdeten Person schwankte die Bereitschaft, lieber sich selber in Gefahr zu bringen zwischen 60% und über 90%. Die Entscheidung wurde hierbei mit ausreichend Zeit zur Abwägung gefordert. Ob eine tatsächliche Entscheidung im Verkehrsraum genau so gefallen wäre, wurde nicht untersucht.

Grundsätzlicher ist die Untersuchung von Wintersberger et al. (2017). Hier wird die Frage gestellt, ob grundsätzlich regelbasiert entschieden werden soll oder zufällig und für unterschiedliche Anwendungsszenarien untersucht, wie nach verschiedenen Kriterien (z.B. Alter der beteiligten Personen) die Entscheidung zur Rettung aussehen könnte. Da sich

Proband\*innen etwa gleich häufig für Regelbasiertheit und Zufall entscheiden, kann hier keine eindeutige Präferenz abgeleitet werden. Wenn aber ein regelbasiertes System eingeführt wird, sprechen sich über 80% der Befragten dafür aus, dass dies für alle Fahrzeuge gleich sein soll, die einzelnen Kriterien also nicht durch Nutzer\*innen oder Eigentümer\*innen individuell konfigurierbar sein sollen.

In einer umfangreichen Studie setzten Sütfield et al. (2017) Proband\*innen in einem Fahrsimulator mit VR-Brille ähnlichen kritischen Entscheidungen aus, bei denen in einer ausweglosen Situation eine Person bzw. Gruppe gerettet und eine andere überfahren werden musste. Auch hier zeigte sich, dass die Entscheidung häufig von verschiedenen persönlichen Kriterien abhängig gemacht wird, etwa Alter oder Geschlecht. Anders als in der zuvor vorgestellten Studie bestand hier ein höherer Zeitdruck für die Entscheidung. Es wurden zwei Varianten getestet, 1 Sekunde und 4 Sekunden, also rein intuitives Verhalten und Verhalten mit sehr kurzem Zeitfenster zur bewussten Entscheidung. Dabei zeigten sich signifikante Unterschiede im Entscheidungsverhalten, dahingehend, dass die Spur im Vergleich zur Ausgangssituation seltener gewechselt wurde, wenn die Zeit etwas länger war, also seltener in das Geschehen eingegriffen wurde. Gleichzeitig zeigte sich auch, dass bei kurzen Entscheidungszeiträumen die Entscheidungen sowohl zwischen Proband\*innen als auch von einer einzelnen Person inkonsistenter sind.

Die Ergebnisse der vorgestellten Quellen zeigen die Notwendigkeit für regelbasierte Systeme, einen moralisch-ethischen Grundkonsens in der Gesellschaft für die relevanten Fahrsituationen zu identifizieren. Methodisch stellt dies eine große Herausforderung dar. In der im folgenden vorgestellten Studie soll hierzu untersucht werden, wie konsistent das Verhalten einer Person ist, wenn sie beliebig Zeit und alle Informationen zur Situation hat im Vergleich zur spontanen Entscheidung, wie sie sich im realen Verkehrsraum darstellt.

### 3 Zielsetzung und Methode

Ein Gedanke, der u.a. von Goodall (2014) aufgebracht wird, ist der Vorschlag, mittels selbstlernender Systeme aus der Beobachtung des Verhaltens von Fahrer\*innen einen Entscheidungsalgorithmus in die Lage zu versetzen, sich in ähnlichen Situationen automatisiert vergleichbar zu verhalten. Dies setzt aber voraus, dass gewünschtes und tatsächliches menschliches Verhalten in Unfallsituationen übereinstimmen, was in dieser Arbeit näher untersucht werden soll. Insbesondere wird dabei untersucht, ob sich das mehrheitlich utilitaristische Verhalten, welches sich u.a. in der im vorangegangenen Abschnitt vorgestellt Arbeiten von Frison et al. (2016) findet, im realen Verhalten widerspiegelt oder der Selbstschutz der Eigentümer\*innen höher wiegt, wie etwa Bonnefon et al. (2015) fanden. Entlang des moralischen Werts von Selbstaufopferung zugunsten anderer Individuen (Sachdeva et al., 2015), der Wirkung sozialer Erwünschtheit (Fisher, 1993) und sog. Zwei-Prozess-Modelle, die automatische (d.h. unbewusste, schnelle, intuitive) Prozesse und kontrollierte (d.h. aufwändige, motivationsabhängige, langsame, durchdacht) Prozesse unterscheiden (Neuberg & Fiske 1987; Kahneman, 2003; Suter & Hertwig, 2011) werden für diese Studie zwei Hypothesen formuliert:

**Hypothese 1:** Proband\*innen treffen in den gleichen Unfallszenarien in der Theorie (d.h. mit ausreichend Zeit zur Überlegung) andere Entscheidungen als im Fahrsimulator (d.h. in einer spontan eintretenden Unfallsituation).

**Hypothese 2:** Die Proband\*innen versuchen sich im Fahrsimulatorversuch selbst zu retten.

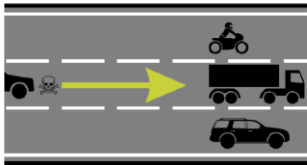
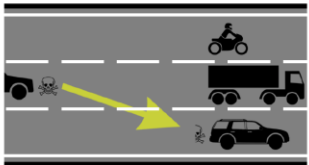
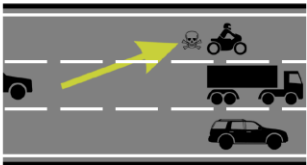
Hierzu wurde ein Versuchsszenario entwickelt. Die verwendete Verkehrssituation sollte so gestaltet sein, dass eine Verkehrsentscheidung sehr offensichtlich einen anderen Menschen mit hoher Wahrscheinlichkeit tötet, die andere eine Lebensgefahr für den/die Fahrer\*in darstellt. Dabei wurde ein Szenario angelehnt an Lin (2015b) umgesetzt. Darin fährt das Fahrzeug auf der mittleren Spur einer dreispurigen Straße mit hoher Geschwindigkeit. Unerwartet taucht vor dem eigenen Fahrzeug ein stehender LKW auf. Ein Abbremsen ist offensichtlich nicht möglich, die Fahrt in den LKW wird mit sehr hoher Wahrscheinlichkeit tödlich enden. Ein Ausweichen auf die linke Spur würde eine\*n Motorradfahrer\*in vermutlich tödlich treffen, ein Ausweichen auf die rechte Spur eine Kollision mit einem SUV bedeuten, was die dessen Insassen verletzen aber nur mit geringer Wahrscheinlichkeit töten würde, für den/die Fahrer\*in aber ein hohes, im Vergleich zum LKW aber geringeres Lebensrisiko bedeuten würde. Um einen Einfluss der rechten bzw. linken Seite zu verhindern, wurde bei der Hälfte der Proband\*innen im Fahrversuch rechts und links gegenüber der obigen Schilderung vertauscht.

Im Folgenden werden Ihnen verschiedene Verkehrsszenarien dargeboten, bei denen es zu Unfällen kommt. Sie haben jeweils drei Entscheidungsmöglichkeiten. Bitte entscheiden Sie möglichst spontan, welche der Optionen am ehesten Ihr Verhalten in dieser Situation abbildet.

**Einleitung Szenarien**

Sie befinden sich mit **130 km/h** auf der **Autobahn**. Der LKW auf der mittleren Spur vor Ihnen bremst abrupt stark ab, eine Kollision mit hoher Geschwindigkeit ist trotz Vollbremsung unvermeidbar. **SZ01**

Sie haben folgende drei Optionen:

1) Sie fahren weiter geradeaus und kollidieren mit dem LKW. Bei diesem Aufprall ist die Wahrscheinlichkeit **sehr hoch**, dass Sie sterben. Der **LKW-Fahrer trägt keinen Personenschaden** davon. Somit würde das Risiko für andere Verkehrsteilnehmer minimiert werden.

2) Sie weichen auf die rechte Spur aus und kollidieren mit dem SUV. Bei diesem Aufprall ist die Wahrscheinlichkeit **hoch**, dass Sie sterben. Der **SUV bietet für die eigenen Passagiere ein hohes Sicherheitslevel**. Das Risiko für einen Personenschaden für die Insassen des SUVs ist jedoch dennoch gegeben.

3) Sie weichen auf die linke Spur aus und kollidieren mit dem Motorrad. Bei diesem Aufprall ist die Wahrscheinlichkeit **hoch**, dass Sie den Unfall **weitestgehend unbeschadet** überstehen, der **Motorradfahrer jedoch stirbt**.

Abbildung 1: Ausschnitt aus dem Fragebogen zum Verhalten in der kritischen Verkehrssituation

Die Studie bestand aus zwei Phasen. Vorab wurden die Proband\*innen zur Teilnahme an einer Online-Befragung (mittel SoSci-Survey) aufgefordert. Nach mindestens einer Woche erfolgte die Fahrsimulator-Studie. In der Online-Befragung wurden demographische Daten, Fahrerfahrung, Risikobereitschaft, Technikaffinität und Kenntnisse zum autonomen Fahren erfasst. Abschließend wurde das Verhalten in verschiedenen Verkehrsszenarien erfragt, darunter die oben geschilderte Situation. Abbildung 1 zeigt den entsprechenden Auszug aus dem Fragebogen.

Mit einem Abstand von mindestens einer Woche wurden die Proband\*innen zur Simulatorstudie eingeladen, um sicherzustellen, dass der Fragebogen nicht mehr in detaillierter Erinnerung war. Die Umsetzung erfolgte in einem statischen Fahrsimulator unter Verwendung der Software SILAB des WIVW. Den Aufbau zeigt Abbildung 2 (links).



Abbildung 2: Links: Fahrsimulator zur Versuchsdurchführung. Rechts: Übernahmeaufforderung

Zu Beginn der Testfahrt erfolgte zunächst eine Einführung in das System und insbesondere in die Aktivierung und Deaktivierung der Fahrautomatik. Vor Eintritt des relevanten Szenarios durchliefen die Proband\*innen zunächst eine kurze Eingewöhnungsphase und längere Fahrt im Automatikmodus. Darin fuhr das Fahrzeug in der mittleren Fahrspur mit 130 km/h. Auf der linken Spur überholten schnellere Fahrzeuge, auf der rechten Spur befanden sich langsamere. Ca. 9 Sekunden vor einer Kollision im Testszenario erfolgte eine Übernahmeaufforderung (visuell und akustisch, siehe Abbildung 2, rechts). Die Zeit reichte aus, um die Kontrolle wieder gut übernehmen zu können (vgl. etwa Gold et al., 2013). Das tatsächliche Verhalten der Testpersonen in der Entscheidungssituation wurde von der Versuchsleitung protokolliert. Um Rechts-/Links-Effekte auszugleichen, wurden LKW und Motorrad bei der Hälfte der Proband\*innen gegenüber der Darstellung in Abbildung 1 vertauscht. Nach Durchführung wurden die Proband\*innen nach einer kurzen Begründung für ihre Entscheidung gefragt.

## 4 Ergebnisse und Interpretation

An der Studie nahmen 12 Proband\*innen teil, 7 männlich und 5 weiblich. Alle Teilnehmer\*innen waren Studierende der Hochschule Ruhr West im Alter zwischen 20 und 32 Jahren ( $M =$

25,4 Jahre) und im Besitz eines PKW-Führerscheins. Acht Proband\*innen gaben an, täglich oder mehrmals pro Woche selbst Auto zu fahren, eine Person gar nicht.

Die Frage nach der eigenen Handlung in der zu untersuchenden Unfallsituation beantworteten sechs Personen mit Option 2 (Auffahren auf den SUV) und jeweils drei Personen mit Option 1 und 3 (Auffahren auf Motorrad bzw. SUV).

Bei acht der zwölf Personen (75%) wich das tatsächliche Verhalten im Fahrsimulator von der Vorabbefragung ab (siehe Tabelle 1). In jeweils zwei Fällen wurde statt zum Motorrad bzw. SUV auszuweichen, wie in der Vorabbefragung angegeben, auf den LKW aufgefahren. Alle vier Begründungen weisen darauf hin, dass die Situation nicht korrekt eingeschätzt wurde. Dies mag zumindest in einigen Fällen der speziellen Testsituation im Fahrsimulator, möglicherweise der fehlenden Fahrdynamik oder einer zu kurzen Eingewöhnungsfahrt geschuldet sein und zeigt die Herausforderung, ein zuverlässiges Versuchsssetup aufzubauen. In anderen Fällen mag es auch darin begründet sein, dass in der Vorabbefragung die einzelnen Überlebenswahrscheinlichkeiten klarer benannt waren, als im Simulatorbild möglicherweise erkennbar war.

Proband*in	1	2	3	4	5	6	7	8	9	10	11	12
Vorabbefragung	SUV	SUV	SUV	MRD	MRD	LKW	LKW	LKW	SUV	MRD	SUV	SUV
Simulator	LKW	SUV	LKW	LKW	LKW	MRD	LKW	LKW	LKW	LKW	SUV	MRD

Tabelle 1: Beabsichtigtes und tatsächliches Verhalten der Proband\*innen in der Konfliktsituation aus Abbildung 1.

Mindestens drei Abweichungen sind inhaltlich begründet. Jeweils ein/e Proband\*in wechselte von der Kollision mit dem LKW bzw. SUV in der Vorabbefragung auf einen Unfall mit dem Motorrad im Fahrsimulatorversuch. Als Begründung wurde eine geringere Gefahr für das eigene Leben angegeben.

Eine weitere Testperson wechselte die Entscheidung von einer Kollision mit dem SUV zum Auffahren auf den LKW mit der Begründung, andere Personen nicht gefährden zu wollen. Die letzte Person mit inkonsistentem Verhalten gab in der Vorabbefragung an, auf die Motorradspur zu wechseln, blieb im Fahrsimulator aber auf der LKW-Spur ohne eine weitere Angabe von Gründen.

Bei drei Testpersonen (davon eine Kollision mit dem SUV und zwei Kollisionen mit dem LKW) war das Verhalten aufgrund der Begründung klar als konsistent einzustufen, andere Personen sollten gar nicht oder möglichst wenig gefährdet werden. Eine weitere Person zeigte das vorhergesagte Verhalten (Ausweichen zum SUV), gab als Begründung einen Reflex an.

Die gewonnenen Ergebnisse unterstützen die Hypothese 1, die Stichprobe ist jedoch für eine verlässliche Aussage zu klein. Zu Hypothese 2 ist nach dieser Studie keine Aussage in eine der beiden Richtungen zu treffen. Zwar deuten zwei begründete Wechsel in diese Richtung, ein anderer widerspricht, insgesamt sind aber deutlich mehr Daten erforderlich, um eine Aussage zu treffen. Die Gesamtzahl der Wechsel legt aber nahe, dies in einer Nachfolgestudie weiter zu untersuchen.

## 5 Fazit und Ausblick

In der Gesamtschau der Ergebnisse ist festzuhalten, dass 75% der Proband\*innen im Fahrsimulator ein anderes Verhalten zeigten, als selbst vorhergesagt. In mindestens 25% der Fälle ist tatsächlich eine bewusste inhaltliche Entscheidung der Grund für den Wechsel, in weiteren 25% war das Handeln konsistent mit ebenfalls konsistenter Begründung. Bei etwa der Hälfte der Proband\*innen ist in diesem Bezug keine eindeutige Zuordnung möglich.

Auch wenn die Stichprobengröße gering und die Auswahl für die Bevölkerung nicht repräsentativ ist, lassen die Ergebnisse die Schlussfolgerung zu, dass a priori geäußerte Verhaltensabsicht und tatsächliches Verhalten nicht in allen Fällen übereinstimmen. Wenn man davon ausgehen kann, dass sich mehrheitlich ethisch gewünschtes Verhalten eher in der Befragung ohne Zeitdruck bei voller Informationsfülle zeigt denn in der tatsächlichen Fahrsituation, so ist dies bei der Konzeption von Systemen zu beachten, die versuchen wollen, aus Fahrer\*innen-Beobachtungen mittels künstlicher Intelligenz ein Verhalten vorzusehen, das den ethischen Maßstäben der Gesellschaft entsprechen soll.

Umgekehrt ist bei der Bewertung von Systemen aber zu berücksichtigen, dass auch Individuen nicht stets diesen Kriterien genügen. Ähnlich wie auch nicht komplette Unfallfreiheit, sondern nur eine geringere Unfallrate als im manuellen Fahren gefordert wird, sollte auch hier „nur“ eine Verbesserung gefordert werden.

Um die Aussagekraft der Studie zu validieren, sollten zunächst am Versuchsaufbau Verbesserungen vorgenommen, um die Immersion zu erhöhen und eine bessere Einschätzung der Situation zu ermöglichen. Dabei kann sowohl eine Verbesserung des Verkehrsszenarios als auch des Hardware Setups in Betracht gezogen werden. Weiterhin sollte der verbesserte Versuch mit einer größeren Stichprobe, die auch besser dem Bevölkerungsquerschnitt entspricht, durchgeführt werden, um statistisch signifikante Daten zu erhalten.

## Literaturverzeichnis

- BMVI (2017). Bundesministerium für Verkehr und digitale Infrastruktur. *Ethik-Kommission Automatisiertes und vernetztes Fahren*, Bericht Juni 2017.
- Bonnefon, J. F., Shariff, A., & Rahwan, I. (2015). *Autonomous vehicles need experimental ethics: Are we ready for utilitarian cars?*. arXiv preprint arXiv:1510.03346.
- Frison, A. K., Wintersberger, P., & Riener, A. (2016). First person trolley problem: Evaluation of drivers' ethical decisions in a driving simulator. In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM. S. 117–122.
- Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20(2), 303–315.

- Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 57, No. 1). Los Angeles, CA: SAGE Publications. S. 1938–1942.
- Goodall, N. (2014). Ethical decision making during automated vehicle crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2424(1), 58–65.
- Kahneman, D. (2003). Perspective on judgement and choice: mapping bounded rationality. *American Psychologist*, 58(9), 697–720.
- Lenz, B., & Fraedrich, E. (2015). Gesellschaftliche und individuelle Akzeptanz des autonomen Fahrens. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Hrsg.), *Autonomes Fahren* (S. 639–660). Berlin: Springer Vieweg.
- Lin, P. (2015a). Why ethics matters for autonomous cars. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Hrsg.), *Autonomes Fahren* (S. 69–85). Berlin: Springer Vieweg.
- Lin, P. (2015b). *The ethical dilemma of self-driving cars*, <https://ed.ted.com/lessons/the-ethical-dilemma-of-self-driving-cars-patrick-lin> (Zugriff: 26.6.2018)
- Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcomes dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, 53(3), 431–444.
- Nyholm, S., & Smids, J. (2016). The ethics of accident-algorithms for self-driving cars: an applied trolley problem? *Ethical Theory and Moral Practice*, 19(5), 1275–1289.
- Rahwan, I., Bonnefon, J. F., & Shariff, A. (2016). *Moral Machine*. Scalable Cooperation, MIT Media Lab. At: <http://moralmachine.mit.edu> (Zugriff: 26.6.2018).
- Sachdeva, S., Iliev, R., Ekhtiari, H., & Dehghani, M. (2015). The Role of Self-Sacrifice in Moral Dilemmas. *PLoS One*, 10(6), e0127409.
- Sütfeld, L. R., Gast, R., König, P., & Pipa, G. (2017). Using virtual reality to assess ethical decisions in road traffic scenarios: applicability of value-of-life-based models and influences of time pressure. *Frontiers in Behavioral Neuroscience*, 11, 122.
- Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition*, 119(3), 454–458.
- Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, 59(2), 204–217.
- Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, 94(6), 1395–1415.
- Wintersberger, P., Frison, A. K., Riener, A., & Thakkar, S. (2017). Do moral robots always fail? Investigating human attitudes towards ethical decisions of automated systems. In *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. S. 1438–1444.