



Responsibility gap or responsibility shift? The attribution of criminal responsibility in human-machine interaction

Monika Simmler

To cite this article: Monika Simmler (2024) Responsibility gap or responsibility shift? The attribution of criminal responsibility in human-machine interaction, *Information, Communication & Society*, 27:6, 1142-1162, DOI: [10.1080/1369118X.2023.2239895](https://doi.org/10.1080/1369118X.2023.2239895)

To link to this article: <https://doi.org/10.1080/1369118X.2023.2239895>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 25 Jul 2023.



[Submit your article to this journal](#)



Article views: 1183



[View related articles](#)



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

Responsibility gap or responsibility shift? The attribution of criminal responsibility in human–machine interaction

Monika Simmler 

Law School, University of St. Gallen, St Gallen, Switzerland

ABSTRACT

The prevalence of automation and advent of intelligent machines have created new constellations in which the attribution of criminal responsibility is complicated. Automation results in complex settings of interaction, while the conduct of technical systems is becoming less determined, predictable, and transparent. There is an ongoing scholarly debate regarding how these developments shape moral and legal agency, as well as the best ways of allocating responsibility. Concerns have been expressed about the emergence of responsibility gaps, calling for ways to ensure accountability. This article explores whether these apprehensions can be empirically substantiated. A factorial survey study ($N = 799$) conducted in Switzerland was used to research attribution mechanisms in criminal cases involving various forms of human-machine interaction. The results revealed that the level of automation significantly affected the attribution of responsibility. In cases of high automation, attribution became considerably more complex and more actors, especially corporate entities, were called to account. A difference in automation level (i.e., the question of how much humans are still ‘in the loop’) had a stronger effect than did the aspect of technology being described as capable of learning. However, the involvement of intelligent technology seemed to have made the responsibility attribution more arbitrary. According to the respondents, there were no discernable responsibility gaps. However, in the present research, significant shifts among the agents called to account were observable. Yet, since these evasion mechanisms are hardly covered by today’s law, there is an ongoing risk of a gap between the desire for punishment and actual legal constructs.

ARTICLE HISTORY

Received 5 December 2022
Accepted 23 June 2023

KEYWORDS

Criminal responsibility;
human–machine interaction;
responsibility gap; agency;
responsibility attribution

1. Introduction

The importance and spread of technical systems have seen unprecedented growth in the last several decades. An increase in technological progress can be observed that goes far beyond anything that existed before. The future social impact of these advances is

CONTACT Monika Simmler  monika.simmler@unisg.ch  University of St. Gallen, Bodanstrasse 3, CH-9000, St Gallen, Switzerland

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

expected to be tremendous (Abbott, 2020). Social change has always had a decisive impact on criminal law, shaping concrete legislation and the principals and fundamentals of criminal responsibility (Killias, 2006). Advanced and even autonomous machines are promising to take over more and more tasks, and thus replace us taking certain actions. Actions are the origin of every criminal liability. When in socio-technical constellations they are delegated to machines and these machines are (partially) autonomous, a challenge is posed to the attribution processes operating under criminal law. Research has focused intensely on the issue of how automation and technical autonomy affect the attribution of criminal responsibility (e.g., Beck, 2016, 2017; Gless et al., 2016; Hallevy, 2010; Simmler & Markwalder, 2019). This debate in criminal legal scholarship is also closely linked to those in other disciplines. The consequences of automation for responsibility and accountability are giving rise to related concerns in the ethics literature, as well as that of other legal disciplines (e.g., Balkin, 2015; Coeckelbergh, 2020; Danaher, 2016; Santosuosso & Bottalico, 2017). The focus is on questions such as the moral and legal agency of machines, who carries what responsibilities for machines, and where the lines run in terms of the use of modern technology. A recurring concern is that of emerging responsibility gaps (Beck, 2016; Matthias, 2004; Santoni de Sio & Mecacci, 2021) or even lawlessness (Balkin, 2015) resulting from the advent of highly complex automation and Artificial Intelligence (AI).

It is evident that the emergence of intelligent technical systems can have an impact on the attribution of responsibility under criminal law. After all, studies have already shown that people hold computers and robots at least partially morally responsible for their actions (Friedman, 1995; Kahn et al., 2012; Kirchkamp & Strobel, 2019). Actual legal assessments will only emerge, however, from concrete court cases, which are as yet still rare. Probably, these rulings will be based on society's perception and interpretation of responsibility with regards to human-machine interactions. Examining these social attribution practices is therefore a fruitful foundation for further scholarly debate. Social perception can provide important (though not binding) indications for future legal assessment.

Accordingly, this research collected and analyzed empirical evidence regarding the question of how automation and technical autonomy affect the assessment of criminal cases, from the perspective of a surveyed student population ($N = 799$). While the impact of technical systems on criminal cases is often referred to in a general fashion, it is important to note that actions cannot be delegated exclusively, or not at all, to machines. It is always a matter of degree. Thus, this study adds to the debate by differentiating among the impacts of different levels and types of automation. In a factorial survey, subjects were confronted with various criminal cases involving different constellations of human-machine interplay. This research presents the findings, with a particular focus on whether concerns regarding a responsibility gap are gaining momentum.

2. Theoretical foundations

2.1. Automation, autonomy, and the distribution of agency

It is important to consider both automation and technical autonomy when evaluating the influence of technology use for agency distribution and, thus, responsibility attribution.

Automation generally refers to a state in which a machine performs tasks or parts of a task formerly executed by humans (Simmler & Frischknecht, 2020). Regarding the defined task, the machine acts independently of a human operator (Nof, 2009). The extent of the control being transferred to the technology is described by the level of automation (Vagia et al., 2016). Thus, automation is not a matter of all or nothing, but rather a ‘continuum of levels’ (Parasuraman et al., 2000).

The attribution of responsibility is based on the attribution of actions. Only those who are capable of acting are regarded as agents and can be (due to this agency) held accountable for their conduct. Socio-technical collaboration consists of at least one human and one technical system. This interaction leads to a distribution of control and, potentially, of agency. The question of the extent to which humans are still ‘in the loop’ with regards to acting and decision-making is crucial and of particular relevance to liability (Santosuosso & Bottalico, 2017), making not only automation *per se* relevant to criminal law, but also the concrete level of automation (Simmler, 2023).

However, the nature of advanced technology is neglected when only differentiating among levels of automation. Advanced or intelligent technical systems are characterized by the growing independence not only from the operator but also from the creator, due to increasing technical autonomy (Simmler, 2023; Simmler & Frischknecht, 2020). The capacity to react to changes in the environment (Alonso & Mondragón, 2004) and ability to interact (Floridi & Sanders, 2004) have been proposed as key factors in technical autonomy. Autonomous technical systems are described as adaptable, as learning artefacts (Floridi & Sanders, 2004; Russell & Norvig, 2014). Technical autonomy is therefore linked to the phenomena of AI and machine learning. Like automation, autonomy is gradually present or absent. Therefore, the level of technical autonomy determines how independent and indeterminate a technical system is in terms of the execution of a task (Simmler & Frischknecht, 2020).

At first glance, it seems apparent that when machines take over acts formerly performed by humans, to the same extent they also assume agency (i.e., the control and power to act). In this sense, distributed tasks lead to distributed agency and responsibility (cf. Rammert, 2012; Strasser, 2022). This delegation of tasks would then excuse the human delegator. Yet, regardless of their level of autonomy, machines have not traditionally been considered responsible subjects. Neither in moral philosophy nor in law they are granted personhood (Simmler & Markwalder, 2019). According to a traditional dualist understanding, it is exclusively humans who act, whereas machines function (Rammert, 2012). It therefore remains unclear how the involvement of (autonomous) technology would be contextualized and to whom the responsibility for their co-action be ascribed. The issue is important, however, especially where this collaboration results in criminal wrongdoing.

2.2. Criminal responsibility of and for machines

Criminal law’s function in society is to stabilize norms through punishment, and such punishment is predicated on placing blame. Acts are considered blameworthy when a subject to whom legal personhood is granted negates criminal norms, even though they could have acted otherwise. Although strict liability offenses are widespread in Anglo-Saxon legal orders (though clearly rejected in most Western European countries)

(Simmler, 2020), *nulla poena sine culpa* represents the essential manifesto of criminal law doctrine. Thus, Anglo-Saxon law requires a ‘guilty mind’ (*mens rea*) and that no defense is present, while for instance in German-speaking countries, one is considered blameworthy only if conduct in accordance with the law was reasonable and the person was mentally competent to behave in that way. However, the meaning of the accusation of guilt is the same in principle and machines are not recognized as being capable of culpability. They are not considered subjects of criminal law (Gless et al., 2016; Simmler & Markwalder, 2019). They are neither blamed nor punished through the legal system. However, the debate over whether AI entities are moral agents is intense and ongoing (Coeckelbergh, 2020; Floridi & Sanders, 2004), and legal scholars have elaborated concepts such as that of electronic personhood (i.e., E-Persons) (Avila Negri, 2021; Beck, 2016). Surprisingly, research has shown that the public does to some extent feel the need to punish artificial artefacts, but they do not consider this punishment to be truly feasible or effective (Lima et al., 2021).

Because machines are not themselves the subjects of criminal law, there is an ongoing debate regarding who (if anyone) is criminally responsible for the behavior of machines (e.g., Beck, 2016, 2017; Gless et al., 2016; Hallevy, 2010; Simmler, 2023; Simmler & Markwalder, 2019). In the case of fully determined and unadvanced technical systems, the answer to this question can usually be found by applying traditional doctrines. However, even then, challenges to ascribing responsibility can emerge. This is because developing, programming, producing, distributing, implementing, and using technical systems usually involves substantial numbers of potentially accountable individuals and companies. In the case of unwanted damage, the question of who may have disregarded their duty of care and acted negligently can seldom be easily answered. The increased complexity and reduced explainability of human-machine interaction in the digital age complicates this considerably (Abbott, 2020). This challenge of complexity is referred to as the problem ‘of many hands’ (Santoni de Sio & Mecacci, 2021), but also as that ‘of many things’ (Coeckelbergh, 2020). This issue is magnified exponentially when adding autonomous technology. Then, not only the responsible subjects must be identified but it furthermore must be established if the machine’s conduct was foreseeable and controllable by them (Beck, 2016). Otherwise, no responsibility can be attributed under criminal law.

The complexity of technology developments and unpredictability of (partially) autonomous machines challenge responsibility attribution (Santoni de Sio & Mecacci, 2021). This is not to say that today’s criminal law doctrines do not generally allow for the attribution of responsibility in cases of human-machine interaction. Intentional acts in which machines are deliberately used to cause harm usually do not cause any issues (Gless et al., 2016; Simmler, 2023). The human who intentionally brings harm is liable to prosecution. Likewise, when an individual or corporation can be identified as not acting in line with an established duty of care, it is called to account under the regime of negligence liability. As said, however, it may often be difficult to identify who lacked compliance. Moreover, it is not always clear what normative expectations are placed upon them regarding not yet established technology. Not every risk-taking is negligent. There is an area of admissible risk accompanying new developments (Beck, 2016); otherwise, innovation would be hindered. It is therefore interesting to evaluate how criminal cases involving human-machine interaction are perceived by social observers. Society is still formulating

expectations regarding such human-machine interactions and moving towards common ground regarding what role criminal law plays in protecting these expectations.

2.3. Responsibility gap

Human-machine interactions may distribute agency, raising the question of whether and what impact this has on the distribution of criminal responsibility. The delegation of acts to machines could serve as an excuse. However, the technical part of a collaboration does not subsume responsibility because machines are not considered legal agents (yet). Other agents that could be called to account for this technical element can only barely be identified (due to the associated complexity), and if identified, cannot always be blamed due to the autonomy (i.e., the unpredictability) of the machine. This is why concerns for emerging responsibility (Chinen, 2016; Matthias, 2004), culpability (Santoni de Sio & Mecacci, 2021), accountability (Mittelstadt et al., 2016), and retribution gaps (Danaher, 2016) have been raised. Due to the aforementioned challenges, such concerns are reasonable.

Conversely, Tigard (2021) has argued that there is no technology-based responsibility gap because moral responsibility is as a 'dynamic and flexible process' that is able to encompass modern technology development. It is indeed plausible that people may simply turn to other agents who have set the machine in motion. Responsibility would then be shifted forward. Either the person who ceded their agency to the machine or the one who helped shape the delegation of action by programming or implementing the technical system would be declared responsible. Alternatively, it may also be that in cases lacking individual responsibility, that responsibility is simply shifted to the corporation or even new agents like E-Persons that have been especially constructed for this purpose (Avila Negri, 2021; Beck, 2016). This would mean that the responsibility gap is closed, or depending on the understanding, never occurred.

Finally, it is imaginable that responsibility gaps occur and are not closed, but merely accepted (Tigard, 2021). Criminal law as *ultima ratio* is not concerned with every economic or societal activity, especially where civil liabilities are regarded as sufficient to guarantee compliance with certain norms. In this sense, the term 'gap' would not be accurate since this would not have to represent a deficiency but instead could be understood as the logical consequence of a changed baseline. Thus, whether a changed attribution of responsibility constitutes a 'gap' also depends on whether one considers such a change to be appropriate. For now, it remains unclear the direction that criminal responsibility in human-machine interaction will take, but learning so is essential if we want to keep criminal law functioning properly in the digital age.

3. Hypotheses

Given the issues described above, this study examined four hypotheses. The latest research has indicated the prevalence of a certain desire to punish machines (Lima et al., 2021), suggesting that technical systems are already being assumed as having some amount of responsibility. However, the law does not currently grant machines agency or personhood. Although the concept of E-Persons has been discussed (e.g., Beck, 2016), it is not expected that such an approach is currently fully reflected in social perception. Thus, it was assumed that when people were asked to evaluate criminal cases

involving human-machine interaction, they would not attribute any criminal responsibility to the machines themselves (H_1).

Automation involves a delegation of tasks, and from a sociological standpoint, is accompanied by a distribution of agency (Rammert, 2012). Such a delegation may simultaneously result in a delegation of responsibility. Therefore, it was further hypothesized that humans acting as users would be attributed less criminal responsibility as the level of automation grew higher (H_2). The assumption of actions by machines would in this case have an excusing effect for the human portion in the interaction. Delegation would lead to exculpation.

It remains unclear to whom responsibility would be attributed with regards to the technical side of human-machine interactions, so long as machines are not regarded as being responsible for themselves. The problem of 'many hands' (Santoni de Sio & Mecacci, 2021), among others, complicates attribution under traditional doctrine. As discussed above, this complication could lead to a responsibility gap (e.g., Chinen, 2016; Matthias, 2004; Mittelstadt et al., 2016); however, this thesis is not or not fully supported by everyone (e.g., Nyholm, 2018; Tigar, 2021). In the present study, it was predicted that the involvement of technology would result in a shift in responsibility away from humans acting situationally towards individuals or corporations in charge of the development and implementation of a technical system (H_3). This hypothesis was formulated because it was expected that people would have the need to identify a responsible actor. Thus, a responsibility gap would not occur, and instead a shifting of responsibility to other agents would be seen.

The final hypothesis was that if the technology involved was (partially) autonomous, the effects described above would intensify (H_4). There is a consensus in the literature that intelligent technology has the potential to challenge criminal responsibility (e.g., Santoni de Sio & Mecacci, 2021). Therefore, it was assumed that technical autonomy would significantly change the way criminal responsibility would be ascribed, as compared to cases of mere automation.

4. Method

4.1. Variables

To test the hypotheses, this study assessed the attribution of criminal responsibility in different cases involving a (potential) criminal offense. As the intention of the study was not to research punitive attitudes but rather to identify the relevance of technical involvement, levels of automation, and technical autonomy, a factorial survey (or vignette study) was chosen (cf. Auspurg & Hinz, 2015). The vignettes described four different cases. They differed experimentally in terms of the independent variable described below, leading to 16 possible variations of case descriptions.

The four vignettes contained a wide variety of potential crime settings and different offenses. Case 1, Railroad, concerned the derailment of a train, resulting in five people dying and 15 others being injured. In this case, the derailment was due either to fault on the part of the train driver or technical system. Case 2, Smart Parking, described the consequences of an incorrectly parked vehicle blocking a fire exit. As a result, firefighters were delayed in reaching the scene, resulting in increased property damage

and two injuries. Depending on the experimental variation, the parking issue was due to the involvement of a smart application that either recommended or automatically executed the improper positioning of the car. In Case 3, Emergency Room, a pregnant woman died because the hospital incorrectly triaged the patient. This triage was executed with varying degrees of technical assistance. In Case 4, Chatbot, a customer was insulted in the customer service chat of a telecom company. The insult originated either from a human being or chatbot acting more or less independently. All cases were described in such a way that none of the actors deliberately wanted to cause harm. They were potentially preventable incidents resulting from the actions or omissions of humans alone or human-machine interactions.

The characteristics (i.e., the ways these criminal cases were described) varied experimentally in four respects (see Table 1). The correspondingly adapted description of the technical involvement in each case served as the independent variable. For Level 1, no technical system was involved in the wrongdoing. For Level 2, the technical system was only involved in a supportive way, with the human being remaining the main agent. For Level 3, the technical system became the essential factor (i.e., it had actual power to influence the situation and the human no longer exercised direct control). For Level 4, the technical system was not only the executing force, but also described as (partially) autonomous, based on its ability to learn. The variation of the vignettes can be exemplified using Case 1. In this case, at Level 1, the train driver was inattentive and thus caused the accident (without technical influence). At Level 2, an autopilot function was involved that reacted incorrectly to a switch being set. The driver should have monitored the autopilot and could have intervened but did not do so in time. In the Level 3 case of high automation, the train derailed because the autopilot was incorrectly programmed and therefore reacted incorrectly to the switch position. The driver could not have intervened. The same applies to Level 4, but there the autopilot was self-learning and further developed on the basis of data it collected itself.

Table 1. Characteristics of the independent variable.

	Variation
Level 1	No automation
Level 2	Low automation
Level 3	High automation
Level 4	Automation and autonomy

The dependent variable was the attribution of criminal responsibility. It was therefore of interest which actors the participants identified as offenders. The respondents were given a wide selection of possible offenders to choose from, and they indicated their choices by ticking a box. This selection consisted of the situationally-acting individual or user of the technical system, company or institution implementing the technical system, programmer (i.e., a human being), software company, and machine itself. In some cases, other participants were also named. Multiple selections were allowed. It was also possible to check a 'Nobody' or 'Other' box.

The control variables were gender and prior criminal law education. In addition, the respondents' affinity for technology was assessed. Following Karrer et al. (2009), five statements were used to rate the options on a Likert scale ranging from 1 'strongly

disagree' to 5 'strongly agree.' These controls were chosen because gender and legal education can affect punitive attitudes and affinity for technology probably has influence on attitudes towards human-machine interactions.

4.2. Data collection

The survey was conducted via a paper questionnaire that was distributed to students at a Swiss university during various lectures given in December 2018. Each of the four randomly composed experimental groups received the same fixed set of four vignettes. Accordingly, the selection was not randomized, but rather systematic (Atzmüller & Steiner, 2010). Each of the four groups assessed one case from each level but never the same case twice.

The relatively large homogeneity and easy accessibility were reasons in favor of selecting students as participants. A homogeneous experimental group was an appropriate sample for the given study design, as this carried a lower risk of unequally distributed third variables (Auspurg & Hinz, 2015). As mentioned, the goal was not to make absolute statements about punitive behavior. Rather, the influence of the independent variable was to be examined as an isolated factor. In this respect, homogeneous experimental groups were advantageous. However, the composition of the sample had to be considered when interpreting the results.

4.3. Limitations

The method chosen was suitable because all respondents were presented with identical case constellations. That would not have been possible if real court judgments were used (Suhling et al., 2005). However, external validity was compromised by the fact that it was a simulation, the cases were simplified compared to real cases, and respondents may have encountered the case differently from how they would have in a real life. Among other simplifications, the analysis did not address the severity or kind of punishment, but only the selection of those to be punished. The analysis would otherwise have been too intricate. It should additionally be noted that the results are not analogous to the behavior of experts. After all, the students were non-specialists. Even if some had undergone basic legal training, they were very likely to judge cases differently from how professional judges would have concluded. Furthermore, the sample is not representative. It could be that other societal groups would perceive such cases and technological involvement differently. An attempt was made to counter this limitation by testing the control variables. Nevertheless, the results cannot be considered to have general validity without further investigation. A final limitation concerns the time of the study. The survey was conducted in late 2018. New technologies are developing rapidly, so it cannot be ruled out that attitudes towards such technologies will also change rapidly and could already be different.

5. Sample

The questionnaire was distributed to a total of 816 students. Seventeen did not complete it. Therefore, the sample consisted of 799 respondents distributed among the four experimental groups ($N_1 = 197$, $N_2 = 201$, $N_3 = 197$, and $N_4 = 204$). Of the total, 58.2% of the

respondents were male and 41.8% female (varying from 55.3% to 63.2% among the groups). The average age was 20.3 years ($N=793$, missing six, varying from 20.1 to 20.4 years among the groups) and ranged from 15 to 57 years. In terms of nationality, 72.3% were Swiss nationals, 13.4% German, 6.3% of other nationalities, 7% with multiple nationalities, and 1% gave no information. For country of origin, 80.3% indicated Switzerland ($N=777$, missing 22). 45.1% of the students studied Business Administration, 17.3% Law, 12.2% Law & Economics, 9.2% International Relations, and 6.4% something else. 9.8% made no statement. 31.6% of participants had already attended a criminal law course ($N=794$, missing 5). The affinity for technology averaged 3.36 ($N=786$, missing 13) on a five-point scale (varying by group between 3.28 and 3.42).

6. Results

6.1. Complexity of responsibility attribution

The study revealed that human-machine interaction led to increased complexity in responsibility attribution. As compared to cases with only humans in control, criminal responsibility was more often placed on not just one but several agents (see Table 2). While 71.7% of respondents in cases of no automation (Level 1) attributed responsibility to only one agent, only 44.7% did so at Level 2, when a technical component was added. This increase in complexity is statistically significant. With high automation and autonomy (Levels 3 and 4), there was no further increase, and rather a slight decrease. In general, the correlation between the level and number of agents called to account was significant ($p < .001$, $r = .107$). The tendency of the complexity of responsibility attribution to increase was observable in all cases. On average, 1.12 responsible agents were selected at Level 1, 1.63 at Level 2, 1.48 at Level 3, and 1.41 at Level 4.

Table 2. Complexity of responsibility attribution: all cases.

Level	N=	Nobody	One agent	Multiple agents
No automation	797	70 (8.8%)*/+	571 (71.7%)***/ooo/+++	156 (19.6%)***/ooo/+++
Low automation	796	49 (6.2%)**/ooo	356 (44.7%)	391 (49.1%)*/ooo
High automation	796	80 (10.1%)	373 (46.9%)	343 (43.1%)
Automation and autonomy	789	96 (12.2%)	386 (48.9%)	307 (38.9%)
Total=	3178	295 (9.3%)	1686 (53.1%)	1197 (37.7%)

Significant difference from the next level: * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

Significant difference from the level after the next one: ° = $p < .05$; °° = $p < .01$; °°° = $p < .001$.

Significant difference between the lowest and highest levels: + = $p < .05$; ++ = $p < .01$; +++ = $p < .001$.

6.2. Responsible agents

6.2.1. Overview

To analyze the results of all cases together, the potential offenders were generalized where possible. This led to results regarding the categories of user, (institutional) implementer, (human) programmer, software company, machine, and a residual category of 'other.' Certain response options were not generalizable. Conversely, not every case involved all categories of potential offenders. These peculiarities were considered when analyzing the results, and especially the percentages. Furthermore, it is of note

that where mentions were of interest in the analysis outlined below, their sum was higher than the sample size N because participants could choose more than one actor (see, e.g., Figure 1).

In sum, the results regarding the agents identified as criminally responsible revealed clear tendencies (see Table 3). As will become more evident below, as automation increased, the user's perceived importance was significantly reduced. At Levels 3 and 4, the implementer of the technical system was very often held accountable. The software company also had surprisingly high relevance in cases of human-machine interaction, while the machine itself was rarely (but still occasionally) blamed. Not all variances from one to another level were significant, as Table 3 shows. However, the difference between low and high automation was particularly remarkable.

Table 3. Responsible agents: all cases.

	No automation ($N = 797$)	Low automation ($N = 796$)	High automation ($N = 796$)	Automation & autonomy ($N = 789$)
Nobody	70 (8.8%)*/+	49 (6.2%)**/ooo	80 (10.1%)	96 (12.2%)
Human actor/user	651 (81.7%)ooo/+++	641 (80.5%)***/ooo	213 (26.8%)	231 (29.3%)
Implementer	176 (22.1%)°ooo/+++	207 (26%)***/ooo	303 (38.1%)	323 (40.9%)
Programmer	–	106 (13.3%)***	183 (23%)***	114 (14.4%)
Software company	–	251 (31.5%)***/ooo	385 (48.4%)*	345 (43.7%)
Machine	–	36 (4.5%)*°	57 (7.2%)	56 (7.1%)
Other	28 (3.5%)	1 (0.1%)	3 (0.4%)	1 (0.1%)

Significant difference from the next level: * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

Significant difference from the level after the next one: ° = $p < .05$; °° = $p < .01$; °°° = $p < .001$.

Significant difference between the lowest and highest levels: + = $p < .05$; ++ = $p < .01$; +++ = $p < .001$.

6.2.2. Responsibility of the user

There was a clear divide between the low and high levels of automation regarding the person acting either independently or as a user of the technical system at the time of offense. Whereas 80.5% (641) of the respondents at Level 2 called the human actor to account, only 26.8% (213) did so at Level 3 (see Table 3). At the same time, the relevance of the other agents increased significantly between Levels 1 and 3 before declining again slightly at Level 4. A clear shift from the human actor to other subjects of responsibility was visible (see Figure 1). Looking

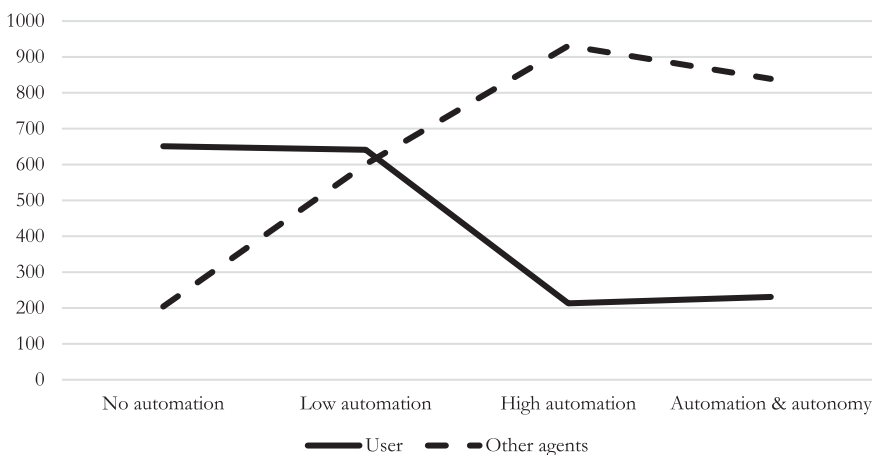


Figure 1. Responsibility of the user as compared to other agents (mentions): all cases.

not at absolute numbers but rather percentages of mentions, the loss of importance of the situational agent becomes even more apparent (see Figure 2). This trend was confirmed by all cases. Interestingly, an effect of the added technical autonomy at Level 4 was hardly visible, although a modest increase in the user's responsibility could be observed. Finally, it should be noted that implementation of low-level automation did not relieve the human actor of responsibility, but rather served to also burden other actors. Here, responsibility actually accumulated (or was diffused).

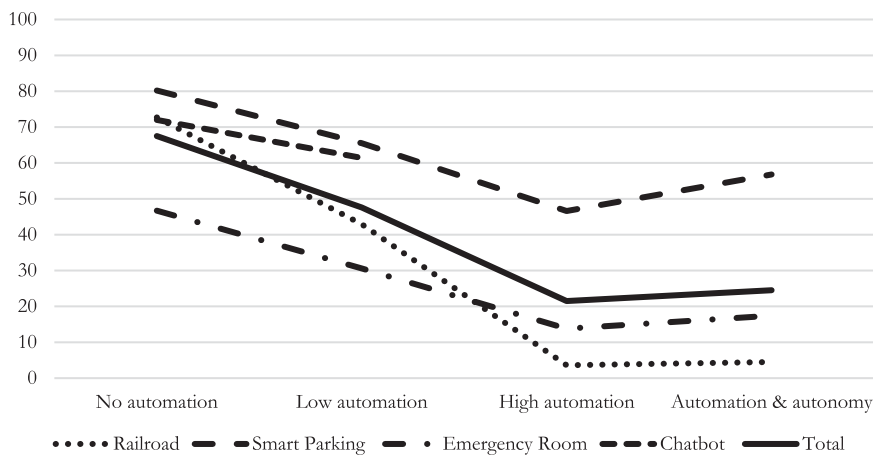


Figure 2. Responsibility of the user (in % of overall mentions): separate cases.

6.2.3. Responsibility of the implementing institution

The frequency with which the implementing company or institution (in Level 1, the employer of the human agent, and in the other levels the one deciding to make use of the technical system) was named by the respondents depended on the level, varying from 22.1% (176) at Level 1 to 40.9% (323) at Level 4 (see Table 3). When looking at the implementer's share of the total number of mentions (see Figure 3), automation initially resulted in a slight decline in relevance, followed by a significant increase in cases of high automation. The decline at Level 2 can be explained by the fact that the total number of mentions was significantly higher here in general, due to the increased complexity of attribution practices. It was rare that the implementing institution was exclusively selected, ranging from 63 at Level 1 to 129 at Level 3 and 130 at Level 4 (i.e., 7.9% to 16.5% of the participants selected the institution alone). While the use of technology at a low automation level relieved these actors, a high automation level placed a burden on them. The relevance of the system's autonomy varied among the cases. In sum, employing a human who acted wrongly or implementing a machine did not *per se* result in very different responsibility attributions. However, once the automation level was high, this circumstance had an incriminating effect for the implementer. In the view of the respondents, high automation led to an increased accountability for institutions in charge of implementation.

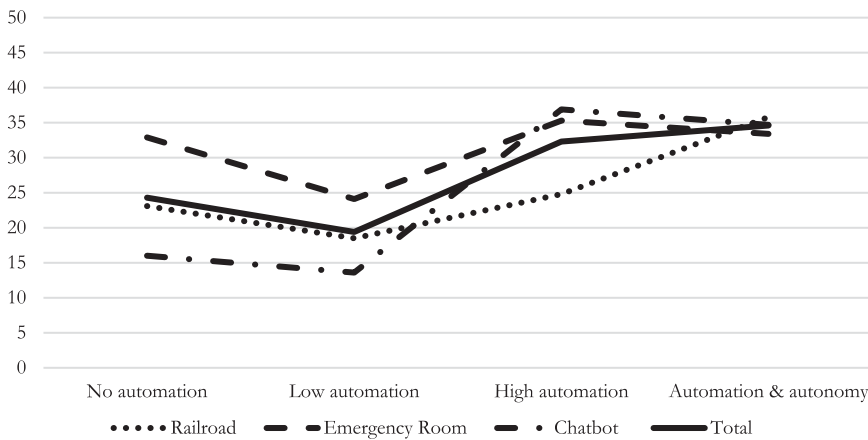


Figure 3. Responsibility of the implementer (in % of overall mentions): separate cases.

6.2.4. Responsibility of the programmer

The programmer (i.e., the human who developed the technical system) was also mentioned regularly, most frequently at the high automation level, where the selection was made by 23% (183) of the respondents (see Table 3). Interestingly, this proportion decreased to 14.4% (114) when technical autonomy was added. Thus, autonomy tended to shift responsibility away from the developers, and to some extent excuse them. This was confirmed when looking at the absolute number of mentions. The programmer received the largest proportion of mentions at Level 3 (see Figure 4). Of course, it is a simplification to assume only a sole human programmer. Nevertheless, tendencies were identifiable regarding how (criminal) responsibility might accompany the development of machines.

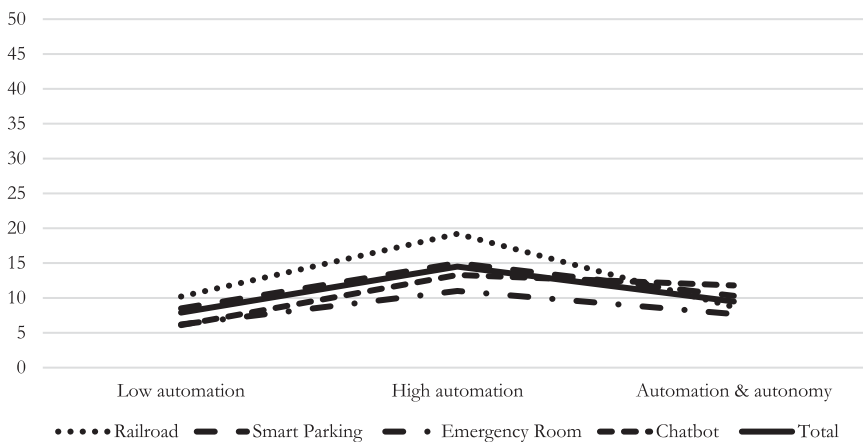


Figure 4. Responsibility of the human programmer (in % of overall mentions): separate cases.

6.2.5. Responsibility of the software company

A similar though less pronounced effect was recognizable for the software company, which was named most frequently by almost half of all respondents (48.4%) at Level 3

and less but still often at Level 4 (43.7%) (see Table 3). The software company was considered substantially responsible in cases of human-machine interaction, particularly at the highest automation level. This impression was confirmed by looking at the proportion of mentions of the software company (see Figure 5). Interestingly, autonomy appeared to have a similar affect regarding the responsibility of the software company to that of the programmer; they seem to be somewhat (a little bit) excused.

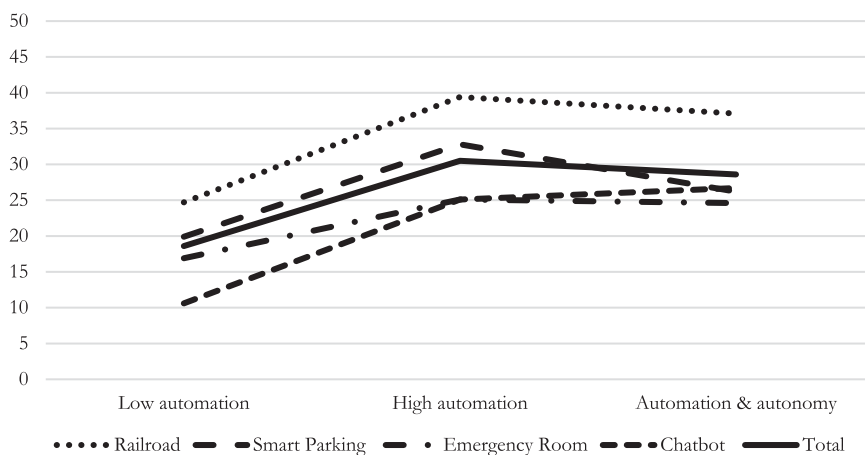


Figure 5. Responsibility of the software company (in % of overall mentions): separate cases.

6.2.6. Responsibility of the machine

The machine itself was mentioned only rarely, by a maximum of 7.2% (57) of the respondents at Level 3 (see Table 3). The shift from low to high automation is statistically significant. Surprisingly, at Level 4, the machine exhibiting enhanced capabilities had no significant effect on the attribution of responsibility to that machine. In this respect, there were differences among the individual cases (see Figure 6).

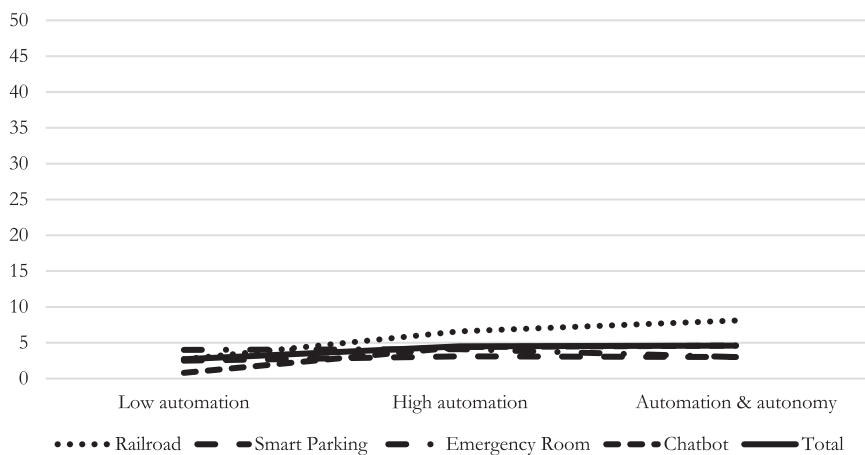


Figure 6. Responsibility of the machine (in % of overall mentions): separate cases.

6.2.7. Waiver of responsibility

The respondents most frequently decided not to hold anyone responsible at Level 4 (12.2%), the case of the use of intelligent technology, while the percentages in other cases ranged from 6.2% at Level 2 to 10.1% at Level 3 (see Table 3). However, there were significant differences among the cases (see Figure 7). In the case of the insulting chatbot, responsibility was most often waived.

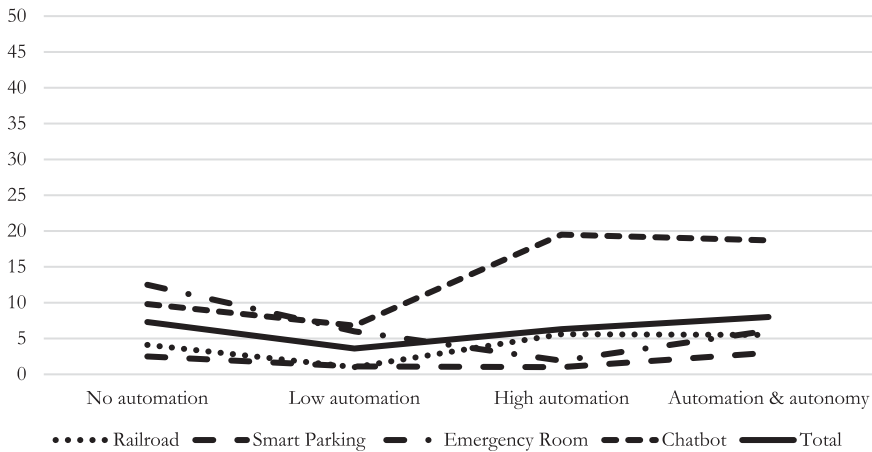


Figure 7. Waiver of attribution of responsibility (in % of overall mentions): separate cases.

6.3. Synthesis: responsibility shift

When considering shifts among the agents (i.e., to whom responsibility was ascribed), an overview makes the increase in complexity very clear (see Figure 8). More actors were attributed responsibility in cases of human-machine interaction. Furthermore, with a higher level of automation, the implementer, programmer, and software company

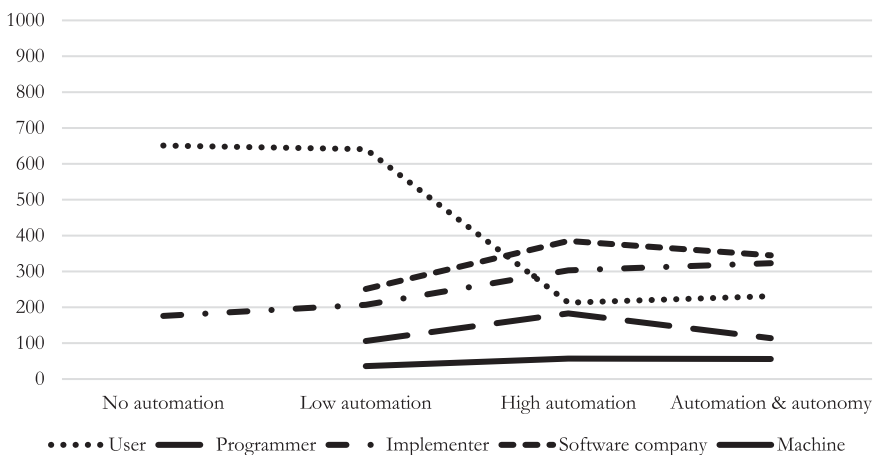


Figure 8. Responsible agents (mentions): all cases.

were increasingly regarded as responsible for the technical part of the collaboration. Further automation led to the human actor ceding not only agency, but also responsibility. However, these data did not allow for a determination of how high the penalty imposed would have been in each case.

Another notable trend was detected from the data. Looking at all the cases individually and in their entirety, corporate criminal liability became significantly more important once there was technical involvement. The respondents often considered the company jointly or solely responsible. As the overview indicated, companies and institutions were named much more often at the highest automation level; up to 72.9% of responses at Level 4 (in the railroad case) called at least one institutional agent to account. Further analysis of corporate responsibility attribution at Levels 2, 3, and 4 found between 464 (34.4%) and 692 (54.8%) mentions referring to a company. Thus, across all cases, there was a clear trend away from individual responsibility and towards corporate responsibility when human-machine interaction affected the case (see [Figure 9](#)). The figures are even more remarkable when one considers that in actual Swiss criminal law, corporate liability plays only an exceedingly minor role.

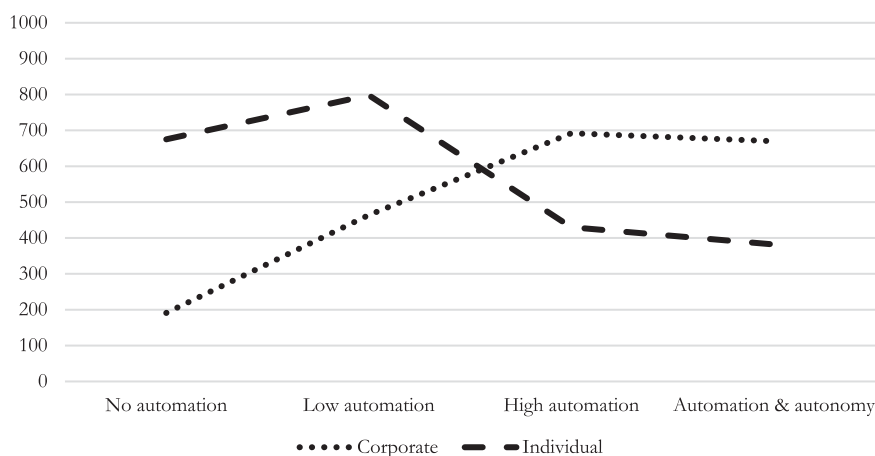


Figure 9. Corporate and individual responsibility (mentions): all cases.

6.4. Controls

As a control variable, the affinity of the participants for technology was measured. The median of the averages of these items was 3.4. The subjects were then divided into two groups (median split); technology savvy (Md above 3.4) and technology-averse (Md below 3.4). Both the technology-savvy and technology-averse groups were comprised of 392 participants (15 missings). When comparing the attribution practices of the two groups, there was only one significant difference. The programmer was called to account more often by the technology-averse group ($p < 0.05$). Apart from that, affinity for technology did not affect the attribution practices.

Controlling for gender resulted in only a few significant differences. At the highest automation level (Level 3), males were slightly more likely to refrain from selecting a

responsible agent than were females ($p < 0.01$). For the implementing organization, differences emerged at the higher automation levels, with males more often penalizing the implementer at Level 3 than did females, while the reverse was true at Level 4 ($p < 0.05$). In the case of the programmer, clearer differences became apparent. Across all levels, females showed a significantly greater willingness to punish the programmer than did males ($p < 0.01$). It remains unclear, however, how this more critical attitude toward the programmer can be explained. It could be speculated that in this regard there is a relationship between the effects of gender and technical affinity. On average, the female participants represented in the sample are less tech-savvy ($M = 2.99$) than the male participants ($M = 3.63$). However, the direction in which these effects manifest themselves and in which way would require further investigation. Finally, females tended to be somewhat significantly more willing to punish the machine at Levels 3 and 4 ($p < 0.01$).

Whether one had attended a criminal law lecture before participating in the study had little effect. People with prior knowledge punished the user more often at Levels 1, 2, and 4 ($p < 0.05$). The programmer was mentioned slightly more often at Levels 2 and 3 ($p < 0.5$). Unsurprisingly, persons somewhat familiar with criminal law named the machine significantly less often at Levels 2 and 4 ($p < 0.5$).

7. Discussion

The first hypothesis (H_1), which stated that no responsibility would be attributed to machines, was not confirmed in absolute terms by this study, though this was only of marginal and negligible relevance. Depending on the constellation and case, between 1% and 12.4% of respondents selected the machine as responsible, of which only 0% to 4.1% attributed sole responsibility. This low percentage can be interpreted to reflect that today machines are not considered criminal subjects. Yet, some people seemed to be open to machine responsibility, an observation that could be relevant in the future, especially assuming that higher levels of autonomy will be part of the equation. Thus, ideas like that of the E-Person (Avila Negri, 2021; Beck, 2016) deserve further thought. For now, most respondents showed no willingness to call the machine to account. Accordingly, the question remains regarding to whom the technical part of the interaction should be attributed (or whether responsibility is waived).

Automation means a cession of situational control to a machine. This ceding of agency is also (to a certain extent) interpreted as a ceding of responsibility, as this study shows. At the same time, automation brings new options regarding the responsible agents upon whom to place blame, making the situation increasingly complicated. These effects are more evident when the level of automation is high. The second hypothesis (H_2), which stated that the more a technical system is involved in the wrongful act, the less criminal responsibility would be attributed to the human actor (user), was consistently confirmed. Those who no longer had control were considered not or less responsible. In this sense, using the technical system resulted in an excusing effect. It was nevertheless remarkable that quite a few respondents continued to blame the human as being responsible, even when the vignette described that they no longer had direct control. Thus, the question arises as to whether someone (and if so, who) is to be held responsible as a substitute. Other agents were indeed named more often, indicating not a waiving but shifting of

criminal responsibility. Likewise, the proposition that distributed agency accompanies distributed responsibility (Coeckelbergh, 2020; Strasser, 2022; Taddeo & Floridi, 2018) was supported. How such distributed responsibility could be captured in future criminal law remains to be seen, and surely also depends on the legal culture of the particular jurisdiction.

If technical systems were involved, evasion mechanisms could be observed, confirming the third hypothesis that this would result in a shift of responsibility towards the programmer and corporation (H_3). This could mean that technology use does not exempt individuals and companies from responsibility, but that they have a duty to construct and use technology carefully. Instead of a responsibility gap, an accumulation or diffusion of responsibility could be observed when technical systems were added. With higher levels of automation, a shift became more visible. At the same time, it is unclear if this shift could be captured by today's criminal law doctrine. The doctrines of corporate responsibility and negligence liability are limited to specific allegations. The mismatch between the human desire to hold someone accountable and the difficulties of responsibility attribution under today's criminal law could result in a retribution gap, as Danaher (2016) has generally predicted.

The evasion mechanisms observed can be interpreted as a way of trying to fill this gap, though they are accompanied by the risk of blaming subjects who did not actually act in a blameworthy fashion. This practice may conflict with the crucial principle of culpability, consequently threatening the effectiveness of criminal law (Simmler, 2020). Respondents did call actors to account, though the description of the case did not indicate an obvious mistake on their part. Participants seemed not to mind the imposition of strict liability. However, there are other possible explanations. After all, it may be reasonable to hold a programmer or software company responsible for a technical failure. They may have made a mistake that could serve as the basis for the accusation. The complicity of the technical system would then not only excuse the user, but also shift responsibility to the events leading to the actual wrongdoing. However, establishing a causal chain and proving that it was not the user who misused the system might remain a challenge. Furthermore, it is conceivable that non-experts do not make clean distinctions between civil and criminal liability.

Another remarkable aspect is that respondents did not refrain from holding corporations responsible. Corporate criminal liability could increase in importance in the digital age. This is an interesting finding, especially for jurisdictions that do not yet accept this concept, usually because of critics of the possibility of establishing a guilty mind in a company and punishing it effectively. In Switzerland, where this study took place, corporate criminal liability can be established only to a very limited extent. It may be that in other jurisdictions (such as the US), the effect made visible in this study would be much more pronounced. Again, it will be important to ensure that companies are not held strictly liable. It must be clear what expectations they have not met. Only then can blame be meaningfully placed and punishment have any stabilizing and preventive effect.

When high automation was paired with technical autonomy, different consequences arose. The complexity of the attribution practice decreased slightly as compared to the level not involving technical autonomy. Still, it remained on a high level. The decrease could be explained by the fact that certain actors may have been released from responsibility when technical autonomy came into play. The programmer and software

company were named less often. Technical autonomy and, thus, a certain unpredictability seem to have had the effect that programmers were attributed less control and consequently somewhat less responsibility. Otherwise, the user was held accountable marginally more often. Likewise, the implementing company was slightly burdened, which could be explained using such an autonomous machine being perceived as a dangerous activity that could *per se* serve as a basis for responsibility. One surprising finding was that technical autonomy did not have a significant effect on the attribution of responsibility to the machine. In sum, the fourth hypothesis (H_4), that technical autonomy magnified the shift in attribution of responsibility, could not be confirmed. Rather, the involvement of autonomous technology led to an increased confusion with regards to attribution practices. Perhaps the assignment also overwhelmed participants, resulting in arbitrary selection. This aspect should be investigated further.

8. Conclusion

Technological and societal change are as interrelated as social and legal transformations. The digital age is shaped by socio-technical interactions being more widespread and involving technology that is becoming increasingly autonomous. Both tendencies can (and probably will) influence criminal law. As the present study reveals, however, this might not happen in the way that is often assumed in the literature. Responsible agents are being sought out and found. There is no responsibility gap (or waiver of responsibility), but rather a shift, at least according to this study's participants. Yet, significant changes in the attribution of responsibility are apparent; automation and autonomy do matter.

In legal practice, two aspects will be challenging when trying to reflect this social sentiment. *First*, corporate criminal liability is not at all (or not very) widespread in many jurisdictions today. However, it could become more prevalent, meaning that specific accusations directed at a company would have to be debated in greater detail. The criminal legal doctrine today is not primarily directed at placing responsibility for technology development and implementation on a company. *Second*, the current dogmatic approach to negligence is likely to reach its limits when dealing with cases of high automation and technical autonomy. This study shows that faulty technology is not simply accepted; a responsible person is found. What exactly they are accused of is not always evident. The duty of care in dealing with technical systems will have to be discussed and a negligence regime perhaps (consciously and carefully) adapted (cf. Beck, 2016). Conversely, care must be taken not to overstretch the scope of the negligence regime. Finally, it must always be kept in mind that criminal liability, unlike civil liability, implies an accusation of personal culpability, leading to punishment: the strongest means the law offers. Criminal law must therefore remain *ultima ratio* and not be used hastily. Nevertheless, this study shows that there is a widespread desire to close gaps in responsibility or rather to prevent gaps from occurring in the first place. At the same time, laws regarding technological advances are developing slowly (Abbott, 2020). The social practices this study revealed must be taken seriously when discussing doctrine and evaluating real-life cases. Therefore, the postulate cannot be limited to the prevention of responsibility gaps. The demand should be to close responsibility gaps where necessary, but to do so in a wise manner. Otherwise, the urge to blame someone

could take over and lead to arbitrary decisions not in line with a criminal law that is functioning properly.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Monika Simmler Assistant Professor of Criminal Law, Law of Criminal Procedure and Criminology.

ORCID

Monika Simmler  <http://orcid.org/0000-0002-2644-8116>

References

- Abbott, R. (2020). *The reasonable robot: Artificial intelligence and the Law*. Cambridge University Press.
- Alonso, E., & Mondragón, E. (2004). Agency, learning and animal-based reinforcement learning. In M. Nickles, M. Rovatsos, & G. Weiss (Eds.), *Agents and computational autonomy: Potential, risks, and solutions* (pp. 1–6). Springer.
- Atzmüller, C., & Steiner, P. M. (2010). Experimental Vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(3), 128–138. <https://doi.org/10.1027/1614-2241/a000014>
- Auspurg, K., & Hinz, T. (2015). *Quantitative applications in the social sciences: Vol. 175. Factorial survey experiments*. Sage.
- Avila Negri, S. M. C. (2021). Robot as legal person: Electronic personhood in robotics and artificial intelligence. *Frontiers in Robotics and AI*, 8, Article 789327. <https://doi.org/10.3389/frobt.2021.789327>
- Balkin, J. M. (2015). The path of robotics Law. *California Law Review Circuit*, 6, 45–60.
- Beck, S. (2016). Intelligent agents and criminal law—Negligence, diffusion of liability and electronic personhood. *Robotics and Autonomous Systems*, 86, 138–143. <https://doi.org/10.1016/j.robot.2016.08.028>
- Beck, S. (2017). Google cars, software agents, autonomous weapons systems – New challenges for criminal Law? In E. Hilgendorf, & U. Seidel (Eds.), *Robotics, autonomics, and the law: Legal issues arising from the AUTONOMICS for industry 4.0 technology programme of the German federal ministry for economic affairs and energy* (pp. 227–252). Nomos.
- Chinen, M. A. (2016). The Co-evolution of autonomous machines and legal responsibility. *Virginia Journal of Law & Technology*, 20, 338–393.
- Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics*, 26(4), 2051–2068. <https://doi.org/10.1007/s11948-019-00146-8>
- Danaher, J. (2016). Robots, law and the retribution gap. *Ethics and Information Technology*, 18(4), 299–309. <https://doi.org/10.1007/s10676-016-9403-3>
- Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, 14(3), 349–379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- Friedman, B. (1995, May 7–11). “It’s the computer’s fault” – reasoning about computers as moral agents [Conference session]. Conference Companion on Human Factors in Computing Systems (CHI 95) (pp. 226–227).

- Gless, S., Silverman, E., & Weigend, T. (2016). If robots cause harm, Who is to blame? Self-driving cars and criminal liability. *New Criminal Law Review*, 19(3), 412–436. <https://doi.org/10.1525/nclr.2016.19.3.412>
- Hallevy, G. (2010). The criminal liability of artificial intelligence entities – from science fiction to legal social control. *Akron Intellectual Property Journal*, 4(2), 171–201.
- Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., Gary, H. E., Reichert, A. L., Freier, N. G., & Severson, R. L. (2012, March 5–8). *Do people hold a humanoid robot morally accountable for the harm it causes?* [Conference session]. 7th ACM/IEEE international conference on human-robot interaction (HRI), Boston, MA, United States.
- Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009). Technikaffinität erfassen – der Fragebogen TA-EG. In A. Lichtenstein, C. Stöbel, & C. Clemens (Eds.), *Der Mensch im Mittelpunkt technischer Systeme. 8. Berliner Werkstatt Mensch-Maschine-Systeme* (pp. 194–201). VDI Verlag GmbH.
- Killias, M. (2006). The Opening and Closing of Breaches – A Theory on Crime Waves, Law Creation and Crime Prevention. *European Journal of Criminology*, 3(1), 11–31. <https://doi.org/10.1177/1477370806059079>
- Kirchkamp, O., & Strobel, C. (2019). Sharing responsibility with a machine. *Journal of Behavioral and Experimental Economics*, 80(3), 25–33. <https://doi.org/10.1016/j.socec.2019.02.010>
- Lima, G., Cha, M., Jeon, C., & Park, K. S. (2021). The conflict between people's urge to punish AI and legal systems. *Frontiers in Robotics and AI*, 8, Article 756242.
- Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6(3), 175–183. <https://doi.org/10.1007/s10676-004-3422-1>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 205395171667967. <https://doi.org/10.1177/2053951716679679>
- Nof, S. Y. (2009). Automation: What it means to US around the world. In S. Y. Nof (Ed.), *Springer handbook of automation* (pp. 13–52). Springer.
- Nyholm, S. (2018). Attributing agency to automated systems: Reflections on human-robot collaborations and responsibility-loci. *Science and Engineering Ethics*, 24(4), 1201–1219. <https://doi.org/10.1007/s11948-017-9943-x>
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3), 286–297. <https://doi.org/10.1109/3468.844354>
- Rammert, W. (2012). Distributed agency and advanced technology. Or: How to analyze constellations of collective inter-agency. In J. Passoth, B. Peuker, & M. Schillmeier (Eds.), *Agency without actors? New approaches to collective action* (pp. 89–112). Routledge.
- Russell, S. J., & Norvig, P. (2014). *Artificial intelligence: A modern approach*. Pearson Education Limited.
- Santoni de Sio, F., & Mecacci, G. (2021). Four responsibility gaps with artificial intelligence: Why they matter and How to address them. *Philosophy & Technology*, 34(4), 1057–1084. <https://doi.org/10.1007/s13347-021-00450-x>
- Santosuosso, A., & Bottalico, B. (2017). Autonomous systems and the law: Why intelligence matters. In E. Hilgendorf, & U. Seidel (Eds.), *Robotics, autonomics, and the law: Legal issues arising from the AUTONOMICS for industry 4.0 technology programme of the German Federal Ministry for economic affairs and energy* (pp. 27–58). Nomos.
- Simmler, M. (2020). Strict liability and the purpose of punishment. *New Criminal Law Review*, 23(4), 516–564. <https://doi.org/10.1525/nclr.2020.23.4.516>
- Simmler, M. (2023). Automation. In P. Caeiro, S. Gless, V. Mitsilegas, M. J. Costa, J. D. Snaijer, & G. Theodorakakou (Eds.), *Elgar encyclopedia of crime and criminal justice*. Elgar Publishing.
- Simmler, M., & Frischknecht, R. (2020). A taxonomy of human-machine collaboration: Capturing automation and technical autonomy. *AI & Society*, 36(1), 239–250. <https://doi.org/10.1007/s00146-020-01004-z>

- Simmler, M., & Markwalder, N. (2019). Guilty robots? – rethinking the nature of culpability and legal personhood in an Age of artificial intelligence. *Criminal Law Forum*, 30(1), 1–31. <https://doi.org/10.1007/s10609-018-9360-0>
- Strasser, A. (2022). Distributed responsibility in human–machine interactions. *AI and Ethics*, 2(3), 523–532. <https://doi.org/10.1007/s43681-021-00109-5>
- Suhling, S., Löbmann, R., & Greve, W. (2005). Zur Messung von Strafeinstellungen: Argumente für den Einsatz von fiktiven Fallgeschichten. *Zeitschrift für Sozialpsychologie*, 36(4), 203–213. <https://doi.org/10.1024/0044-3514.36.4.203>
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good: An ethical framework will help to harness the potential of AI while keeping humans in control. *Science*, 361(6404), 751–752. <https://doi.org/10.1126/science.aat5991>
- Tigard, D. W. (2021). There is no techno-responsibility gap. *Philosophy & Technology*, 34(3), 589–607. <https://doi.org/10.1007/s13347-020-00414-7>
- Vagia, M., Transeth, A. A., & Fjerdingen, S. A. (2016). A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? *Applied Ergonomics*, 53(1), 190–202. <https://doi.org/10.1016/j.apergo.2015.09.013>