



Levels of Driving Automation from a User's Perspective: How Are the Levels Represented in the User's Mental Model?

Hendrik Homans^(✉), Jonas Radlmayr, and Klaus Bengler

Chair of Ergonomics, Technical University of Munich, Boltzmannstraße 15,
85747 Munich, Garching, Germany
{hendrik.homans, jonas.radlmayr, bengler}@tum.de

Abstract. In this explorative study, we focused on the number of levels of driving automation as part of the mental model, using a newly created online questionnaire and analyzing results using Principal Component Analysis. The online questionnaire consisted of 20 automated driving functions that were described in short sentences. The 247 participants subjectively rated the degree of automation of the 20 functions on a 7-point rating scale. Using Principal Component Analysis, the ratings were summarized into groups represented by the components based on correlative relationships. The results yielded three components, which can be interpreted as three levels of driving automation. These three levels differ from the well-established taxonomies, showing that users do not differentiate between five or six levels of driving automation.

Keywords: Mental models · Levels of automation · Vehicle automation · Human-automation interaction · Analysis and evaluation · Autonomous driving

1 Introduction

Automated driving has recently become a big topic in society, research and development. From a human factors point of view, a lot of research is done on topics such as driver attention [1] or take-over performance [2]. All these studies and publications refer to comparable taxonomies of automation from different organizations like SAE [3], NHTSA [4] or the German Federal Highway Research Institute BASt [5]. This work takes a critical view of these levels of driving automation and their perception with the general driver. To describe and distinguish the capabilities of automated driving functions, several organizations created levels of driving automation that follow a similar classification of automation in vehicles. The most common taxonomies refer to either 5 or 6 levels of driving automation and range from “Driver Only” or “No Automation” to “Driverless” or “Full Automation.” The higher the degree of automation, the greater the part of the driving task that is taken over by the technical system [6]. Mental models reflect the individual understanding of a user interacting with a system: which parts the system consists of, how the system works, and why it works [7]. If mental models are wrong or incomplete, wrong actions may consequently be chosen, which may result in unsafe or inefficient use and operation [6]. Based on a

mental model of automated systems, users can derive the fundamental allocation of control and responsibility between humans and machines [6]. This helps the user to derive the functional and responsibility limits of the automated systems. The driver needs to be aware of these limits to enable a safe interaction with the system [8]. There are various methods to measure or assess mental models, such as questionnaires, interviews, behavioral observations and card sorting. If the mental models of the users are unknown, there is the probability of a mismatch to the real world, which can be associated with different risks. This effect is well known from human computer interaction and was discussed by [9]. Original Equipment Manufacturers (OEMs) communicate their system dissimilarly, e.g. comparable systems are called “Autopilot” [10] or “Traffic Jam Assist” [11]. In comparison to other domains that establish automation (aviation, process control), there is no dedicated process of teaching and training, which could lead to an inconsistent model between the developer, organization and user. Based on the literature and the well-established taxonomies, this work addresses the following research questions: How are the levels of automation represented from the user’s perspective?

2 Method

The users rated 20 automated driving functions subjectively on a scale regarding the degree of automation in an online questionnaire. With the Principal Component Analysis (PCA), these driving functions can be divided into groups based on correlative relationships. The method can be used to reduce a larger set of measured variables or items to a smaller group of underlying dimensions [12], called components or factors. With the PCA, the rated driving functions can be summarized into groups represented by the components. Therefore, it is suitable for this research question: in which and how many groups can the functions be summarized based on their degree of automation? At the end of the survey period, 280 complete datasets were gathered. Four records were excluded due to missing preconditions. In addition, 29 records were excluded from the analysis due to unrealistically quick completion. This was based on system-generated values from *Soscisurvey*. The final sample consisted of $n = 247$ participants. The sample consisted of $n = 135$ male and $n = 108$ female participants, with 4 participants not providing their gender. The mean age was $m = 32.71$ years (range: 19–74 years, $SD = 10.72$). All participants had a valid driver’s license and German as their native language. The theoretical knowledge of automated driving functions, in general, is balanced: 41% stated they had “moderate” knowledge; 23% had “low” and 23% had “high” knowledge. Practical experience was less present in the sample compared to theoretical knowledge: 42% only had “low” experience with automated driving functions; 23% had “none” and 21% “moderate” practical experience. Almost 14% in total had “high” or “very high” experience. The 20 driving functions were described in short sentences. Every driving function was described by name and a brief description focused on functionality and constraints, for example:

Cruise Control: The driver sets a speed that is kept by the car. The driver must steer and, if necessary, adjust the speed and brake.

Robot Taxi: Allows for autonomous driving from start to finish - without restrictions: a driver is not necessary.

The names of the 20 driving functions in the order they were presented in the questionnaire were: 1. Robot Taxi ($M = 6.87$, $SD = 0.57$), 2. Night Vision ($M = 1.81$, $SD = 1.15$), 3. Cruise Control ($M = 5.28$, $SD = 1.12$), 4. Collision Protection ($M = 4.18$, $SD = 1.28$), 5. Chauffeur ($M = 5.28$, $SD = 1.09$), 6. Evasive Assistant ($M = 4.32$, $SD = 1.22$), 7. Extended Chauffeur ($M = 5.89$, $SD = 0.97$), 8. Lane Departure Warning ($M = 1.94$, $SD = 1.01$), 9. Emergency Stop Assist ($M = 4.67$, $SD = 1.52$), 10. Traffic Jam Assistant ($M = 4.51$, $SD = 1.20$), 11. Lane Keeping Assist ($M = 3.67$, $SD = 1.23$), 12. Collision Warning ($M = 1.96$, $SD = 1.00$), 13. Traffic Jam Pilot ($M = 5.62$, $SD = 1.22$), 14. Autobahn/Highway Pilot ($M = 6.18$, $SD = 0.96$), 15. Shift Support ($M = 1.50$, $SD = 0.91$), 16. Park Steering Assist ($M = 3.54$, $SD = 1.21$), 17. Adaptive Cruise Control ($M = 3.56$, $SD = 1.22$), 18. Remote Parking, ($M = 5.43$, $SD = 1.39$) 19. Adaptive Cruise Control with Steering Assistance ($M = 4.57$, $SD = 1.20$), 20. Parking Garage Pilot ($M = 6.49$, $SD = 1.00$). The 20 driving functions in the questionnaire were rated on a 7-point rating scale. All 20 functions cover all driving situations of automated driving and the full range from manual driving to autonomous driving and contained already implemented and future functions. This was ensured by experts researching this topic.

For the statistical calculation it was important that the items in an online questionnaire cover the whole spectrum of the rating scale [13]: “no automation” and “full automation.” In an expert discussion, it was ensured that the complete range of driving functions was evenly represented within the 20 functions. The scale was presented as an optical numeric rating scale with 7 levels (unipolar). The number of scale points is chosen based on the literature and the results of the pre-tests. [14] concluded that a 7-point scale corresponds to people's natural judgement, which is why it is preferable to all other scales. The driving functions in the questionnaire were not grouped according to the levels of driving automation or functional areas but are randomly arranged as far as possible. The two pre-tests were conducted with a total of 41 participants to uncover possible weaknesses in the formulation, sequence, length and design of the questionnaire and the items. The online questionnaire was implemented on the platform www.soscisurvey.de and the survey period lasted six weeks. The questionnaire was mainly distributed via social networks and email. In addition to the online questionnaire, 11 questionnaires were completed in printed form.

As a precondition for participation, a valid driver's license and German as their native language was required on the first website to avoid possible language barriers. Both preconditions were verified in the demographic data. The participants could take part in a raffle for online shopping vouchers: 2×50 Euro and 4×25 Euro. Anonymity was ensured by all participants via the system *Soscisurvey*.

Statistical evaluations were performed in SPSS 21 (IBM). The descriptive evaluation shows that the lowest mean of level of automation are the functions “Shift Support”, “Night Vision” and “Lane Departure Warning”. The highest ratings of the degree of automation are the functions “Robot Taxi”, “Parking Garage Pilot” and “Highway Pilot”. The largest standard deviation exists in the function “Emergency Stop Assist”.

With 0.86, the Kaiser–Meyer–Olkin measure of sampling adequacy (MSA) is “meritorious” [15]. Bartlett’s test of sphericity was significant ($\chi^2(190) = 2403$, $p < .001$). Mean communalities over 20 functions are $M = 0.62$ (Range: 0.31–0.74). The lowest communalities can be found for “Robot Taxi” with $h^2 = 0.31$ and “Cruise Control” with $h^2 = 0.36$. There should be at least three variables per factor for the factor extraction [13]. Factor loadings < 0.3 can be neglected [13]. A principal component analysis was performed with 20 variables using a varimax rotation. The varimax rotation was chosen to simplify the interpretation of the components. The extraction was based on substantive considerations and interpretations of the factor loadings as well as on the Scree-Test according to [16] and the Kaiser-Guttman criterion [17]. The rotated solution shows a relatively unique structure in which all major components have high factor loadings. There are no high transverse loads, which is why many functions only load on one main component, which simplifies a clear assignment of the functions to one group. Exceptions are the following functions:

- The “Adaptive Cruise Control” function heavily loads on components 1 (.546) and 2 (.542)
- The “Lane Keeping Assist” function loads on components 1 (.369), 2 (.470) and 3 (.467)
- The “Robot Taxi” loads negatively only on component 1 but with $-.516$. It can also be seen in the correlation matrix that this function does not correlate positively higher with any other function than with $r = .19$ (Motorway Pilot). In addition, this function has the lowest commonality of $h^2 = 0.31$, so suitability for the PCA should be considered critical.
- The “Chauffeur” was taken out of the statistical calculation because two functions were named very similarly: “Chauffeur” and “Extended Chauffeur”. Analysis with both functions within one PCA showed one component with only these two functions. A PCA with only one of the “Chauffeur”-functions showed more perspicuous results. Further analysis was conducted leaving out each one of the “Chauffeur”-functions: both PCAs showed a very similar result with similar statistical values for both “Chauffeur”-functions.

The analysis shows a 3-component solution. Based on the online questionnaire and the principal component analysis, a mental model of levels of driving automation could be established. It was possible to extract a clear 3-component solution based on the used driving functions. The components contain the following driving functions with their factor loadings in brackets.

Component 1 (Eigenvalue: 6.625): None to low automation: Lane Departure Warning System (.821), Shift Support (.821), Night Vision (.788), Collision Warning (.767), Cruise Control (.565), Adaptive Cruise Control (.546).

Component 2 (Eigenvalue: 3.22): Medium automation: Evasive Assistant (.807), Extended Chauffeur (.604), Collision Protection (.569), Emergency Stop Assist (.561), Lane Keeping Assist (.479).

Component 3 (Eigenvalue 1.169): High to very high automation: Remote Parking (.799), Parking Garage Pilot (.793), Traffic Jam Pilot (.672), Autobahn/Highway Pilot (.663), Adaptive Cruise Control with Steering Assistance (.660), Park Steering Assist (.535), Traffic Jam Assistant (.507).

3 Discussion and Limitations

The results on the mental models of automation levels are not in line with the taxonomies that are used by OEMs and the community of experts. This suggests that automated driving functions are not sophisticatedly distinguished between 5 or 6 levels from a current user's point of view. The China Industry Innovation Alliance for Intelligent and Connected Vehicles [18] also utilizes a taxonomy with three groups of automated and connected driving, supporting the need for a new understanding and critical view of the established taxonomies. If users are to understand and interact with automated driving functions in a safe and acceptable way, their mental model should match the functions they interact with in the future. If the results are compared to the existing taxonomies, from a user's point of view, the differences between functions are considered in less detail (Table 1).

Table 1. Comparison between taxonomies and the results of our study

	Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
NHTSA [4]	No automation	Function-specific automation	Combined function automation	Limited-self-driving automation	Full self-driving automation	
SAE [3]	No automation	Assisted	Partial automation	Conditional automation	High automation	Full automation
Mental model	None to low automation		Medium automation		High to very high automation	

It is understandable that technical experts rely on another understanding and view in comparison to non-expert users (see [9]). For example, in research and functional development, technical possibilities and limits need to be identified, which may require detailed technical and functional taxonomies. However, especially in complex and safety-critical systems, development and design should be user-centered. The different understanding and mental models between experts and users present a problem when users get in touch with the products made by these experts without focusing on the user's mental models. Here, the HMI of automated driving functions and the advertisement of such functions plays a predominant role to avoid wrong expectations, misuse and frustration. A stronger orientation in the development and introduction of automated systems, especially considering existing mental models, will increase acceptance and system trust.

The automated driving functions were only described textually. The correct understanding of the functions cannot be guaranteed. One option would be to let the participants experience the functions themselves in simulators or real vehicles (see for example [8]). The names of the functions could, regardless of their function, provide hints to the classification; such as the words "Assistant" and "Pilot." For example, all driving functions that include the word "Help" or "Warning" are summarized in the first component. Also, the functions with the addition "Pilot" can be found in one component. However, it would be conceivable that these functions would be assessed

differently by the participants if they had a different (or no) name. On the one hand, it is unclear whether the 3-level model is also demonstrably present in the population. On the other hand, the study was conducted in Germany in German, which is the reason why the results are based on German terms. It would be reasonable to replicate this study in other language and cultural areas. In addition, mental models differ between different people: there is not only one mental model. They change and vary depending on expertise and experience. It would be conceivable to collect and compare the mental models of non-experts and experts with a sufficiently large sample and to quantify distances within and between groups.

References

1. Radlmayr, J., Brüch, K., Schmidt, K., Solbeck, C., Wehner, T.: Peripheral monitoring of traffic in conditionally automated driving. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **62**, 1828–1832 (2018)
2. Gold, C., Körber, M., Lechner, D., Bengler, K.: Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density. *Hum. Factors* **58**, 642–652 (2016)
3. Society of Automotive Engineers: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. (SAE Standard J3016_201609). PA: SAE International, Warrendale (2016)
4. National Highway Traffic Safety Administration: Preliminary Statement of Policy Concerning Automated Vehicles. <http://www.nhtsa.gov/staticfiles/rulemaking/pdf/>
5. Gasser, T.M. (ed.): Rechtsfolgen zunehmender Fahrzeugautomatisierung; Gemeinsamer Schlussbericht der Projektgruppe; Bericht zum Forschungsprojekt F (Legal consequences of increasing vehicle automation; joint final report of the project group; Report on the research project F). *Wirtschaftsverl. NW Verl. für neue Wiss, Bremerhaven* (2012)
6. Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., Beller, J.: Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations. *Cogn. Technol. Work* **14**, 3–18 (2012)
7. Rouse, W.B., Morris, N.M.: On looking into the black box: prospects and limits in the search for mental models. *Psychol. Bull.* **100**, 349–363 (1986)
8. Beggiano, M., Krems, J.F.: The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transp. Res. Part F: Traffic Psychol. Behav.* **18**, 47–57 (2013)
9. Norman, D.: *The design of everyday things*. Basic Books, New York, NY (2013)
10. Tesla, I.: *Autopilot: Full Self-Driving Hardware*. <https://www.tesla.com/autopilot>
11. Volkswagen AG: *Traffic Jam Assist*. <http://www.volkswagen.co.uk/technology/proximity-sensing/traffic-jam-assist>
12. Field, A.: *Discovering Statistics Using IBM SPSS Statistics and Sex and Drugs and Rock ‘n’ Roll*. Sage, Los Angeles, London, New Delhi, Singapore, Washington DC (2013)
13. Bühner, M.: *Einführung in die Test- und Fragebogenkonstruktion (Introduction to Test Design and Questionnaire Design)*. Pearson Studium, München (2006)
14. Krosnick, J.A.: Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl. Cognit. Psychol.* **5**, 213–236 (1991)
15. Kaiser, H.F.: An index of factorial simplicity. *Psychometrika* **39**, 31–36 (1974)

16. Cattell, R.B.: The scree test for the number of factors. *Multivar. Behav. Res.* **1**, 245–276 (1966)
17. Kaiser, H.F.: The application of electronic computers to factor analysis. *Educ. Psychol. Measur.* **20**, 141–151 (1960)
18. Li, K.: The Base Platform of ICV System and Its Industrialization Approach. SIP-adus Workshop; Cross-Ministerial Strategic Innovation Promotion Program, Council for Science, Government of Japan. Tokyo, Japan (2018)