

Article

# Mode Awareness and Automated Driving—What Is It and How Can It Be Measured?

Christina Kurpiers \*, Bianca Biebl \*, Julia Mejia Hernandez and Florian Raisch

BMW Group, Knorrstrasse 147, 80788 München, Germany; Julia.Mejia-Hernandez@bmw.de (J.M.H.); Florian.Raisch@bmw.de (F.R.)

\* Correspondence: christina.kurpiers@bmw.de (C.K.); Bianca.Biebl@bmw.de (B.B.)

Received: 27 April 2020; Accepted: 18 May 2020; Published: 21 May 2020



**Abstract:** In SAE (Society of Automotive Engineers) Level 2, the driver has to monitor the traffic situation and system performance at all times, whereas the system assumes responsibility within a certain operational design domain in SAE Level 3. The different responsibility allocation in these automation modes requires the driver to always be aware of the currently active system and its limits to ensure a safe drive. For that reason, current research focuses on identifying factors that might promote mode awareness. There is, however, no gold standard for measuring mode awareness and different approaches are used to assess this highly complex construct. This circumstance complicates the comparability and validity of study results. We thus propose a measurement method that combines the knowledge and the behavior pillar of mode awareness. The latter is represented by the relational attention ratio in manual, Level 2 and Level 3 driving as well as the controllability of a system limit in Level 2. The knowledge aspect of mode awareness is operationalized by a questionnaire on the mental model for the automation systems after an initial instruction as well as an extensive enquiry following the driving sequence. Further assessments of system trust, engagement in non-driving related tasks and subjective mode awareness are proposed.

**Keywords:** mode awareness; measurement method; automated driving; SAE Level 2; SAE Level 3

## 1. The Relevance of Automation

Within the next few years, technical advances will enable the development of vehicles that can transport users to their destination without human input. This exclusion of drivers from the control and guidance tasks eliminates human errors and as a result leads to increased road safety [1,2]. The technical complexity of such systems however does not allow a direct switch from manual to fully autonomous driving. Consequently, various car manufacturers are currently developing semi-autonomous systems that can manage some but not all driving functions. The level of automation in these systems is called partially automated driving (PAD; Level 2) according to the taxonomy by SAE International [3]. PAD systems can control longitudinal and lateral acceleration. Nevertheless, these systems require constant monitoring of their performance, traffic and the surrounding by the driver. In contrast to fully autonomous driving, PAD is still prone to human errors like inattention or distraction since the driver has the role of a supervisory controller who acts in collaboration with the system. This automation system cannot detect all its limits and errors, which is why the driver is responsible for intervening if necessary even without a preceding warning or take-over request [4]. This responsibility allocation will change with the introduction of conditionally automated driving (CAD; Level 3). These systems also control longitudinal and lateral acceleration and thus resemble PAD. Contrary to Level 2 however, CAD can detect all system limits itself and will request the user to take over within a certain time frame if necessary. As such, the driver is not required to be attentive to the system's status or traffic when CAD is active and he or she is then allowed to engage in non-driving related tasks (NDRT). Taken

together, the main difference between Level 2 and Level 3 systems is the driver's responsibility for the driving task and the concomitant obligation to pay attention to the traffic situation in PAD but not CAD.

The safety of assisted driving functions is therefore reliant on the user's awareness of the currently active system and the knowledge about his or her responsibilities in this automated driving mode. This understanding is naturally aggravated if PAD and CAD are available within the same vehicle and if both systems are repeatedly activated within one drive [5,6]. It is especially safety critical if the user neglects the monitoring task during PAD because he or she might not notice the system reaching a limit in time. The danger for such an improper behavior is especially increased if the system works perfectly because the user might not expect any system limits [7,8]. In conclusion, it is of great importance to secure a good understanding and clear differentiation of the responsibilities in PAD and CAD. Various measures are currently being developed and tested to provide a so called mode awareness like the issuance of attention requests [9], hands On/Off options [10], the inclusion of one or multiple automation modes within a drive [5] and manual drives in-between periods of automated driving [11]. To assess the effect of such measures on mode awareness, it is however necessary to define mode awareness and to develop appropriate measurement methods first. This article aims to give an overview on the concept of mode awareness and to present a newly developed approach to measure mode awareness during alternating manual, PAD and CAD drives.

## 2. Constructs Concerning Monitoring Behavior

Before addressing the measurement of mode awareness, it is important to define this complex construct first and differentiate it from otherwise variables. The following chapter will provide an overview over all relevant constructs to mode awareness.

### 2.1. Situation Awareness

Mode awareness is similar but not identical to the concept of situation awareness. The latter is constituted of sufficient knowledge about the vehicle's surrounding, the current state of the automation, the system's task performance and the driver's own tasks and responsibilities [12]. If the driver lacks situation awareness, critical situations might be identified too late so that the driver cannot take compensating actions to resolve the situation [12]. According to Endsley and Kiris [13] the extent of situation awareness depends on three factors: automation information presentation; vigilance, monitoring and trust; engagement. Since systems with a higher reliability of autonomy go along with less attention on traffic and system performance, situation awareness is often reduced in higher automation levels [14]. That is why the driver should be given a sufficient amount of take-over time in order to get back in-the-loop before driving manually [15].

### 2.2. Mode Awareness

According to [16], there are two kinds of mode awareness: the awareness of the existence of different automation levels and the awareness of the currently active mode. While both aspects are necessary for mode compliant behavior, the latter is at particular risk when a vehicle incorporates two or more automation levels and is intended to be the focus of this paper [6]. Mode awareness is a subconstruct of situation awareness that merely excludes the knowledge about the current situation and surrounding [17]. It comprises the knowledge about the currently active automation system, its performance level and the driver's tasks and responsibilities [6]. Similar to situation awareness, mode awareness is established by the perception and correct interpretation of system information, the build-up of knowledge and finally the prediction of future system behavior [6,17,18]. A deficit can arise on any of these levels. Most common are however a misinterpretation of the systems' behavior and symbols (mode confusion) or a lack of knowledge about the systems (mental model).

### 2.3. Mode Confusion and Mode Errors

Mode confusion is one possible reason for deficient mode awareness [17,19]. It can be described as a kind of automation surprise, where the system does not behave according to the user's expectations. In the case of mode confusion, the user loses track of which system is currently active or what kind of behavior is appropriate for which mode. Mode confusion is safety critical [20] because it can lead to mode errors. This term describes behavior that fits the assumed but not the actual active automation level [6,21]. It results from an erroneous combination of information in the mental model [22]. Mode confusion can arise if a driver experiences two or more systems when changing between vehicles or when multiple systems are available within one vehicle, of which the latter represents a greater risk for mode confusion. The likelihood for mode confusion further increases if the systems appear similar for the user, e.g., in the case of PAD and CAD [6]. As a result, drivers might engage in NDRTs while driving in PAD and thus neglect their monitoring task. This can be highly dangerous if the system reaches its limit without the driver noticing, which can lead to collisions.

### 2.4. Mental Model

An awareness of the currently active automation mode itself is not sufficient for the creation of mode awareness. In addition, the user must have a correct mental model concerning the automation systems. Mental models are internal representations of a system that are formed by interacting with the system. These models do not need to contain correct technical details as long as users understand the functional characteristics of the system. Mental models can differ greatly in complexity depending on existing knowledge about the system, experience from interacting with the system and education [18,23].

### 2.5. Overtrust

As mentioned earlier, mode awareness is essential for a correct amount of monitoring and the controllability of system limits. Even if users have adequate declarative knowledge about the currently active system, its function and the users' own responsibility, they might however not behave according to the requirements of the automated system [24]. Next to fatigue, risk tolerance, boredom or extrinsic motivation, the greatest danger for such an improper behavior is an inappropriate level of trust in the system. In general, trust describes the attitude of users to let a system support them in situations characterized by uncertainty and potential danger [25]. This trust influences the usage of the automated system. In case of under-trust, users will tend to disuse the system because of the subjectively increased work load and risk [26]. This state should be avoided because a disuse of the automated system will decrease the customer value of the vehicle [27]. Van Loon and Martens [28] for example describe three factors that might be positively impacted by increased automated highway systems: a reduction of traffic congestions, a more economic driving style with a concomitant conservation of resources as well as increased traffic safety. According to the authors, the latter is especially improved after users get used to the new technologies or the system assumes increasing parts of the driving tasks.

In respect of safety in use, the more pertinent problem is over-trust. This blind trust in a seemingly perfect system can result in a misuse of the system beyond its functional limits and thus in a safety risk. In the case of PAD, users will presumably show a decreased monitoring of the driving scene and the system performance with an increased trust in automation because they assume that everything will be working properly [29,30]. This so called complacency is especially provoked if participants simultaneously have to perform multiple tasks which reduces the amount of cognitive resources available for monitoring [31]. Ironically, over-trust and its adjunctive misuse is elicited by a highly reliably functioning system because users will hardly ever experience the system limits they theoretically know about [8,32].

### 3. Measurement of Mode Awareness

The development of a measurement method for mode awareness is crucial for the serial implementation of automated systems since mode inappropriate behavior increases the risk for critical take-over scenarios and crashes. Principally, there are multiple methods to examine mode awareness. It is, however, very difficult to identify a technique which allows a measurement of all subjective and objective aspects of mode awareness. Various potential approaches will be presented and discussed in the following chapter.

#### 3.1. Subjective Measurement Methods

Surely, the simplest way of getting insight into the user's mode awareness is by simply asking the driver via self-rating scales or interviews (e.g., [33,34]). Both methods give fast and explicit information about the user's state and can be used directly after a use case of interest or subsequent to the entire driving sequence. As with any subjective measurement method, it is however subject to a personal bias. Self-ratings on the user's assessment of his or her mode awareness are furthermore insufficient, because users might not understand the complexity of this construct in its entirety. Since misconceptions in the mental model can inherently not be detected by the users themselves, it is not advisable to use self-ratings as an indicator for mode awareness. An interview meets some of these flaws by allowing a standardized, and thus, partly objective assessment of mode awareness. In order to cover all aspects of mode unawareness however you need to identify all potential problems beforehand, which is not only time-consuming but also improbable. Additionally, interviews present multiple difficulties concerning study designs. If they are conducted while driving, the cognitive distraction might confound the driving performance. The conduction of multiple interviews (e.g., before and after an experimental manipulation) also poses the risk of influencing the mental model because the mere reproduction of information increases the knowledge level [35]. If the interview is conducted subsequent to the drive to avoid these confounding factors, the time interval between the driving scenario and the interview may lead to memory distortions, which in turn reduces the validity of information. One method to counteract the disadvantages of self-rating scales and interviews but maintain explicit information about the user's inner processes is a driver commentary (e.g., [36]). This method does not have the problem of memory loss or the need for predefined mode awareness deficits, because it aims to gather all thoughts of a user directly while using the system. The greatest benefit of this method is surely its flexibility towards individual and situational differences. The lack of standardization on the other hand complicates quantitative and comparative analyses. Furthermore, it might lower the ecological validity of a study because the simple instruction to formulate all thoughts might change these thoughts and interfere with the driving task.

#### 3.2. Objective Measurement Methods

Another approach to investigate mode awareness is the use of objective measurements, which eliminates all subjective distortions and focuses on the actual user behavior. As illustrated previously, the main difference between PAD and CAD is the allocation of responsibility between the system and the driver and as such the required amount of monitoring [3,37]. Monitoring implies the placing of visual attention to the street or control instruments, which is most often accompanied by a corresponding eye movement [38]. Therefore, gaze behavior counts as a good indicator for mode awareness. To interpret gaze behavior in terms of mode awareness however you need a comparison value like a drive in another automation level. Another indicator of the user's knowledge about his or her responsibility is the interaction with NDRTs. The engagement in tasks like e.g., smartphone apps, in-vehicle information system, phoning or eating [39] will reduce the time of gaze spent on the traffic or system functionality. Consequently, it covers similar aspects concerning gaze behavior as mode awareness but is more restricted to the engagement in specific tasks. It must be noted that an engagement in NDRTs cannot always be implemented because of the study design or legal requirements.

Ultimately, the main interest in mode awareness does not lie in monitoring behavior, distractions or declarative knowledge itself, but in the consequential driving performance. This includes the take-over performance and the handling of critical situations, like e.g., the reaction time (time until gaze redirects from the NDRT to the road or control instruments; time until hands on the steering wheel and time until the first take-over reaction is performed), the time-to-collision (TTC), the maximum lateral and longitudinal acceleration and crash rate among others [40]. Possible take-over situations can range from uncritical switches between automation levels to undetected system failures in PAD (e.g., following a tar track instead of the line marking). The controllability of such situations is of special interest because of its safety implication. It must however be noted that driver behavior in such take-over situations is not specific to mode awareness and cannot indicate mode awareness problems on its own. It might, for example, be influenced by momentary inattention, fatigue, the familiarity with take-over situations or the individual participant's driving skills. It can thus only be interpreted against the backdrop of the attention ratio during the drive, the pre-existing knowledge and the post-enquiry.

That is why some studies (e.g., [41,42]) look for certain behavior patterns that are likely to be specific to mode unawareness. Mode confusion for example could become apparent when the system reaches a system limit. In addition, a user might grab the steering wheel during CAD, press random buttons repeatedly or show facial cues of confusion. These behavioral characteristics can however vary between participants and might not always occur during a drive, which is why their comparability is reduced. Furthermore, this behavior cannot be ascribed to mode awareness for sure without a follow-up interview. A user might for example put the hands on the steering wheel for comfort or by habit and not because of a misunderstanding of the currently active automation mode.

### 3.3. Combination of Measurement Methods

Mode awareness is a complex construct that circumferences sufficient knowledge about the system and its limits as well as behavioral aspects while driving. Currently, there is no gold standard for measuring all aspects of mode awareness, which is why most authors use a combination of multiple methods. Victor et al. [43], for instance, examined mode awareness during a PAD drive and added a take-over situation at the end of the drive because of an obstacle on the road. Mode awareness was operationalized by subjective as well as objective variables. The former consisted of a questionnaire on trust and open interview questions on impulse to intervene as well as the realization of the need to intervene. Objective data comprised response process variables and glance variables. Generally, a mixture of subjective and objective indicators is advisable for a valid interpretation of the declarative knowledge of users about the current system and its functionality as well as the behavior according to system requirements. In this case, the behavior aspect is assessed sufficiently by analyzing gaze behavior as well as take-over performance and the handling of critical situations. While the short interviews after the drive can give an impression on trust level and situation awareness, however they do not allow an evaluation of the user's mental model and mode confusion.

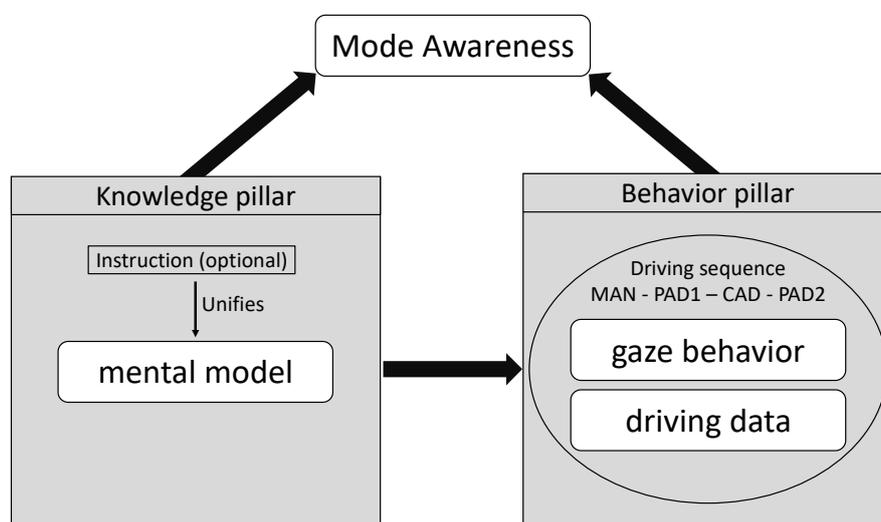
Another approach to measuring mode awareness can be found in a study by Wang and Söffker [44]. The authors investigated six driving scenarios. The implementation of both Level 2 and Three in these scenarios is reasonable for studying mode awareness in a worst case approach [5]. Mode awareness was operationalized by a situation awareness questionnaire but the authors also measured take-over time and quality in case of system failures and the engagement in NDRTs. While these measures do provide subjective and objective information, the lack of monitoring data does not allow a full objective interpretation of mode awareness. Furthermore, it has to be noted that the mid-questionnaire between the scenarios only included six questions on mode awareness, which is very little for such a complex construct.

Othersen [18] conducted a study on situation and mode awareness. Objective measures consisted of driving parameters, specifically reaction and take-over time, the quality of reactions and potential deactivating of the system. Furthermore, the author examined gaze behavior and video data as well as the performance in an audio-verbal NDRT. Subjective data circumferenced items on mode

confusion, monitoring behavior, responsibilities during the drive and critical situations as well as the user's take-over performance. This approach covers many aspects of mode awareness like knowledge about the user's monitoring task (objective and subjective), awareness of the currently active mode (subjective) as well as the resulting take-over performance. The analysis of gaze behavior was however conducted absolutely without systematically comparing different drives to a baseline. In addition, the closed self-rating scale does not provide detailed information about the user's responsibilities. On the contrary, a comprehensive assessment of the user's mental model is crucial to define the cause of a potential lack of monitoring behavior.

#### 4. A Subjective and Objective Measurement Method for Mode Awareness

In order to assess all aspects of mode awareness, we wanted to develop a new method that combines subjective and objective information in a worst case scenario. This approach allows the assessment of all major aspects of mode awareness (see Figure 1): the knowledge about which mode is currently active and the knowledge about the system's abilities and limits (knowledge pillar) as well as the resulting mode compliant behavior (behavior pillar).



**Figure 1.** Mode awareness can be subdivided into a knowledge and a behavior pillar, which are measured separately in our proposed study design. All white frames represent dependent variables.

##### 4.1. Knowledge Pillar

One aspect of the definition of mode awareness according to [6] is sufficient knowledge about the system. Therefore, an assessment of mode awareness should include the measurement of the participant's mental model, which should be conducted before the experimental drive. In studies questioning the effectiveness of certain methods to promote mode awareness, it is vital to first instruct the participants on the automation systems because a different amount of preknowledge can impact the effectiveness of such methods. The subsequent knowledge test can thus ensure a homogenous level of existing knowledge before the drive. Such an extensive instruction is, however, not advisable when studies aim to hedge mode awareness in order to get approval for an automated system. The initial knowledge test then serves as a first indicator of mode awareness during the drive.

It is, furthermore, important to include an extensive post-enquiry to test the amount of knowledge after the drive. That allows a conclusion on the knowledge on the systems' limits, human-machine interface (HMI) and the driver's responsibilities during the drive. This second conductance of the test is furthermore relevant in order to measure the change of the mental model due to a driving sequence or (if applicable) certain experimental manipulation.

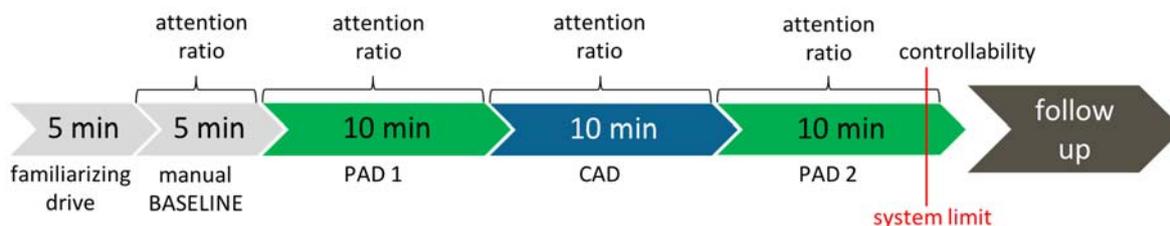
The questionnaire for the mental model we developed consists of five parts. At the beginning, participants are asked to subjectively rate their knowledge about all assistance systems of interest on a 7-point Likert scale. This subjective rating is followed by an objective evaluation of the user’s knowledge. They are first asked to formulate the two main aspects of each system. These statements are then evaluated by the examiner on the basis of a rating system, which categorizes information in mandatory and optional information. This is followed by detailed questions on various aspects of the assistance system, mainly the systems’ limits and abilities as well as the responsibility of the driver. These statements have to be assigned to the respective assistance system or alternatively classified as true or false. Lastly, the participants are tested on their knowledge about the HMI and the handling of the systems’ (de-)activation. That mainly functions as an indicator for mode confusion in the drive since insecurities about the corresponding icons for each mode as well as the correct button for activating and deactivating the systems can easily lead to confusion about the currently active system.

#### 4.2. Behavior Pillar

##### 4.2.1. Design

Declarative knowledge about the systems and their capabilities is necessary but not sufficient for mode awareness. Drivers might for instance technically be well aware of the currently active system and his or her responsibilities but still neglect the monitoring task because of over-trust [29]. A distracted or inattentive driver might then not notice a PAD system reaching its limit and thus crash. This use case is just one potential scenario and certainly represents a worst case setting. In order to ensure the safety of driver assistance systems in studies however a worst case approach is necessary [45].

We propose the following study design to validly measure mode compliant behavior in a worst case scenario (see Figure 2).



**Figure 2.** A schematic depiction of the driving sequences. A first familiarizing drive and a manual baseline are followed by a sequence of driving partially automated driving (PAD), conditionally automated driving (CAD) and then PAD again. Mode awareness is operationalized by the comparison of attention to driving related areas of interest during the drives and the controllability of a system limit at the end of the second PAD drive.

The drive starts subsequent to the theoretical instruction and the questionnaire on the mental model with a familiarizing drive. This drive is crucial to eliminate the influence of prior experiences with driver assistance systems, the make of the car or potential situational factors (e.g., being in a driving simulator). Depending on the research question, the familiarizing drive can contain short drives in all assistance modes including the switches between them or just a manual drive. It is advisable to then start a short period of driving manually as a baseline. The gaze behavior and driving data during this drive serve as comparison values for all subsequent automated drives. As mentioned earlier, a frequent switch between these automation modes is especially challenging for maintaining mode awareness, since the functions seem very similar for the user. In line with a worst case approach, we thus recommend including multiple switches between automation modes within the study. Between CAD and PAD, the latter has fewer situational requirements, which is why a first switch from manual driving to PAD is the most ecologically valid option. It also allows the assessment of the first contact of drivers with PAD as a baseline value. After a certain period of driving PAD, the system should

then enable the switch to CAD. This drive should be terminated by a take-over request to initiate the last driving sequence in PAD. Until this point, both Level 2 and Level 3 would have worked perfectly without reaching any unexpected system limits. This is in line with a worst case approach since the high performance level makes it difficult to distinguish between both systems. By definition however, PAD systems might reach their limit without giving a warning or take-over request, e.g., because they accidentally follow a tar track instead of the actual lane. It is advisable to include such a scenario to assess the controllability of a potentially critical situation. That is crucial for driving safety and actually of higher importance than monitoring behavior. We recommend a silent system error at the end of the second PAD drive, by driving straight ahead instead of following the curved road. Without intervention of the driver the vehicle would then crash with the crash barrier or drive on the adjacent patch of grass.

The time frame of these drives can be chosen according to resources and research question. Generally, a longer time-frame will lead to more reliable data. A longer time-frame will however lead to increased driver fatigue [46]. Multiple internal studies showed that participants need 5 to 10 min to get used to the system. To avoid the influence of fatigue but ensure a sufficient amount of data we thus advise a duration of approximately 8 to 10 min per automated drive. Studies by Kurpiers et al. [9] and Feldhütter et al. [47] confirmed the assumption that this is an appropriate time frame to avoid insecurities when handling the system while simultaneously avoiding fatigue. Certain research questions and participant characteristics might however require the adaptation of these time slots. It must also be noted that the study design proposed in Figure 2 is only applicable in this form for studies in driving simulators. The uncontrollability of on-road studies may not allow the strict adherence to the proposed time frames because of interchanging road conditions and environments. This study design is however an appropriate basis for measuring mode awareness in a simulated environment. As such, it serves as a good tool to test changes in the automated function during development to ensure their security. Furthermore, it can be used to check the effectiveness of measures to increase mode awareness.

#### 4.2.2. Attention Ratio

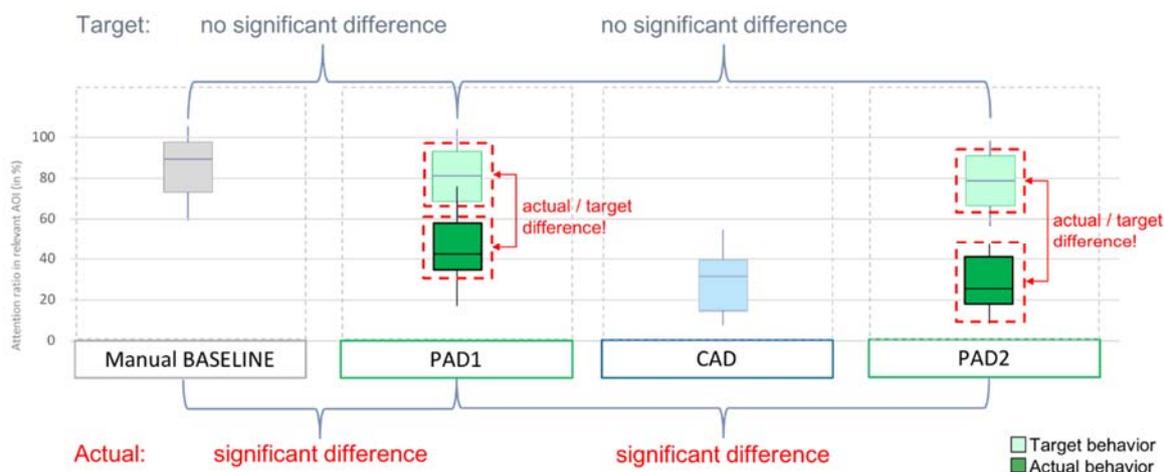
The aim of the manual-PAD-CAD-PAD sequence is to assess the participants' behavior in respect to the mode dependent responsibilities for the user. The requirement to keep the attention on the traffic at all times in PAD but not CAD is surely the greatest difference between these two automation systems and the monitoring behavior therefore a suitable operationalization for mode awareness. As most shifts in visual attention go along with a shift of gaze, glance behavior can be used for operationalization of mode awareness [48]. The most interesting metric of glance behavior is the attention ratio, which represents the percentage of time that a participant's gaze is directed in a certain area of interest (AOI) in relation to the total duration of each driving phase. One AOI of particular interest is the road center slightly below the horizon [49,50], since hazard and event detection requires focal vision [51]. Further driving relevant areas within the visual field are the instrument cluster to monitor the system's status as well as the lanes to the left and right, the side mirrors and the rearview mirror for traffic monitoring. If NDRTs are used (e.g., smartphone apps or games in the central information display), their location should be evaluated as a relevant non-driving related AOI to track the attention ratio to the NDRT.

#### 4.2.3. Target and Actual Values

When PAD is active, the driver is assisted in the lateral and longitudinal guidance of the vehicle. Similar to manual driving, all other responsibility lies with the driver [3]. Since the driver must be able to detect all system limits and take-over at all times even without warning during PAD, the attention ratio to the traffic situation and the system's performance should not differ from that in manual driving. A mode aware driver should furthermore not show any significant differences in gaze behavior between the first and the second PAD drive despite the interposed CAD sequence. During CAD on the other side, it is expected to find a reduced amount of monitoring behavior compared to manual and PAD

driving, since the driver is allowed and instructed to engage in NDRTs [3]. While the amount of monitoring in PAD is safety critical, a comparison between CAD and a manual or a PAD drive can mainly serve as an indicator for the quality of discrimination concerning the user’s tasks. Lack of such a discrepancy in monitoring behavior is however not necessarily evidence for mode unawareness since users might also deliberately want to monitor the CAD function and their surroundings.

This proposed gaze behavior has been tested with similar designs in various studies [9,51]. In the static simulator study by Feldhütter et al. [47] for example, participants showed an attention ratio to road center of 89% during manual driving, which was significantly reduced to 51% and 18% in the first and second PAD drive respectively with a significant decrease from the first to the second PAD drive. This is a characteristic example for a mode awareness deficit, since the monitoring task was neglected during the PAD drives compared to manual driving, which was intensified by the CAD drive in between (schematic depiction in Figure 3).



**Figure 3.** A schematic depiction of the target and actual gaze behavior during manual, PAD and CAD driving. Mode awareness can be assumed if there is no significant difference between the attention ratio in the manual and the first PAD drive or a significant decline in monitoring behavior from the first to the second PAD drive.

#### 4.2.4. Controllability

Next to the monitoring behavior itself, the safety of automated vehicles in Level 2 and Level 3 is highly dependent on the user’s ability to manage system limits. In the proposed worst case scenario of a silent system error in the second PAD drive, the car will keep driving straight ahead while the track makes a curve. The controllability of this situation can be assessed by the ability of the driver to keep the car on track. Potential parameters are the amount of the vehicle in surface area that has crossed the track before the driver intervenes and the crash rate. Feldhütter et al. [47] for example found that only 16% of participants intervened before the car had left the track and 29% did not take-over before the car had left the track completely. This clearly demonstrates the dangers of insufficient monitoring behavior during PAD that results from a deficit in mode awareness. It has to be noted however, that a bad performance in the take-over scenario cannot necessarily be ascribed to a lack of mode awareness. As a result, we advise a qualitative interrogation on the take-over scenario after the drive.

#### 4.3. Additional Variables

The main problem when using gaze data is its lack of specificity since it can be influenced by factors like extrinsic motivation or boredom [24], risk tolerance [11], a faulty mental model and mode confusion [6] or over-trust [29]. Questionnaires before and after the drive are thus essential to ascribe a lack of monitoring and controllability to a concrete source. Next to the before mentioned knowledge test, one important assessment is the evaluation of trust in the automated system, (e.g., the automation

trust scale (ATS) by Jian, Bisantz, and Drury [52]; the questionnaire on human-computer trust by Madsen and Gregor [53]), because deficits in mode awareness and over-trust cannot be distinguished without background information on the user's experience and mindset. In addition, a subjective test for mode awareness (like the one used by Othersen et al. [18]) might be useful in many cases. Furthermore, it is of great value to add various questions, e.g., concerning the perception of the events during the silent system error, the engagement in NDRTs during PAD, a lack of engagement in NDRTs during CAD, automation surprises and other subjective data. The specific choice of questions should be based on the individual characteristics of the driving behavior and the examiner's observations.

## 5. Limitations and Benefits

Despite this approach's theoretical soundness, the validity of the proposed study design has not been calculated yet. The data in [9,47] that result from study designs using the proposed method can give a first impression of its applicability but allows no testimony on the validity of the measurement approach. The main reason for this lies in the circumstance that there is no best practice for measuring mode awareness that could be compared to the results of the suggested approach. Furthermore, the interpretation of mode awareness in our study design is based on a number of different variables that need to be encountered as a whole. As a mixture of quantitative and qualitative measure, it is hardly possible to calculate one parameter for mode awareness that might be used in a validation process. In addition, this design for measuring mode awareness is not applicable in all study designs. First of all, the addition of a critical take-over scenario at the end of the second PAD drive is obviously impossible in on-road car studies. The only solution to evaluate controllability is to look for naturally arising system limits of PAD and assess the take-over quality of the participants. Any study in real traffic is, furthermore, liable to uncontrollable circumstances like weather, traffic and road works that might influence the availability of the assistance systems. Second, some research questions might call for a variation of drives compared to the proposed design, which might change the values of mode awareness. Furthermore, since the order of drives is essential to the assessment of mode awareness, it is not advisable to alter the sequence of the automated drives. That way, however, the attention ratio in the second PAD drive might already be reduced because of tiredness or exhaustion. That should be factored in by performing an objective sleepiness rating, e.g., the Karolinska sleepiness scale (KSS; [54]).

When testing the human-machine interaction of automated functions, the aim of most studies is to predict user behavior in the field. In order to secure the safety of the function, it is important to prove the robustness of the function even in worst cases. Wickens [45] actually argue that accidents in aviation are often caused by worst-case performers in worst-case situations. That is why extreme cases should not be treated as outliers in a normal distribution but considered for safety issues. Consequently, the proposed study design is an appropriate approach to testing mode awareness in PAD and CAD. In addition, this design is eligible for measuring mode awareness in different scenarios because the switch from PAD to CAD and back to PAD allows the relative comparison of attention ratio in Level 2. The use of absolute values on the other hand would lead to misinterpretations, since attention ratio itself will differ greatly between a simulator study without any actual danger and a real car study on public highways.

## 6. Conclusions

We propose a study design to assess mode awareness by focusing on its behavioral aspect, more precisely the attention ratio while driving in PAD, CAD and then PAD again in addition to the controllability of a critical take-over scenario at the end of the second PAD drive. Questionnaires and interviews on the mental mode, trust, the engagement in NDRTs and other observations during the drive will enable the examiner to extract the source of a potential negligence of the monitoring behavior during PAD. Taken together, we feel positive about the potential of this approach to cover all aspects of mode awareness while differentiating it from similar constructs. Further validation of the proposed design and assessment technique is required for a further evaluation.

**Author Contributions:** Conceptualization, C.K. and B.B.; Investigation, C.K.; Methodology, C.K., B.B. and J.M.H.; Project administration, C.K.; Writing—original draft, C.K. and B.B.; Writing—review & editing, C.K., B.B., J.M.H. and F.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** We thank the Chair of Ergonomics of the Technical University Munich for the collaboration. Especially, we would like to thank Klaus Bengler, Anna Feldhütter, Moritz Körber and Michael Rettenmaier for providing the static driving simulator of the Technical University Munich and for conducting three studies that investigated different measures for maintaining mode awareness with the developed measurement method.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fagnant, D.J.; Kockelman, K. Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transp. Res. Part A Pol. Practice* **2015**, *77*, 167–181. [[CrossRef](#)]
2. Treat, J.; Tumbas, N.; McDonald, S.; Shinar, D.; Hume, R. *Tri-Level Study of the Causes of Traffic Accidents. Executive Summary*; U.S. Department of Transportation: Washington, DC, USA, 1979; pp. 1–328.
3. SAE International. *Taxonomy and Definitions for Terms Related to Driving Automation Systems for on-Road Motor Vehicles*; SAE International: Warrendale, PA, USA, 2018.
4. Endsley, M.R. From here to autonomy: Lessons learned from human-automation research. *Hum. Factors* **2017**, *59*, 5–27. [[CrossRef](#)] [[PubMed](#)]
5. Feldhütter, A.; Segler, C.; Bengler, K. Does shifting between conditionally and partially automated driving lead to a loss of mode awareness? In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Los Angeles, CA, USA, 17–21 July 2017; Springer International Publishing: Basel, Switzerland, 2018. [[CrossRef](#)]
6. Sarter, N.B.; Woods, D.D. How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Hum. Factors* **1995**, *37*, 5–19. [[CrossRef](#)]
7. Strauch, B. Ironies of automation: Still unresolved after all these years. *IEEE T. Hum.-Mach. Syst.* **2017**, *48*, 419–433. [[CrossRef](#)]
8. Bainbridge, L. Ironies of automation. *IFAC Proc. Vol.* **1983**, *15*, 129–135. [[CrossRef](#)]
9. Kurpiers, C.; Lechner, D.; Raisch, F. The influence of a gaze direction based attention request to maintain mode awareness. In Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles, Eindhoven, The Netherlands, 10–13 June 2019.
10. Naujoks, F.; Purucker, C.; Neukum, A.; Wolter, S.; Steiger, R. Controllability of partially automated driving functions—Does it matter whether drivers are allowed to take their hands off the steering wheel? *Transp. Res. Part A Traff. Psych. Beh.* **2015**, *35*, 185–198. [[CrossRef](#)]
11. Boos, A. Habituation Effects in Automated Driving: The Influences of Exposure Duration and Exposure Frequency on Visual Monitoring Behaviour. Master’s Thesis, Technical University Munich, Munich, Germany, 2018.
12. Endsley, M.R. Toward a theory of situation awareness in dynamic systems. *Hum. Factors* **1995**, *37*, 32–64. [[CrossRef](#)]
13. Endsley, M.R.; Kiris, E.O. The out-of-the-loop performance problem and level of control in automation. *Hum. Factors* **1995**, *37*, 381–394. [[CrossRef](#)]
14. Kaber, D.B.; Endsley, M.R. The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theor. Issues Ergon.* **2004**, *5*, 113–153. [[CrossRef](#)]
15. Gold, C.; Damböck, D.; Lorenz, L.; Bengler, K. “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, San Diego, CA, USA, 30 September–4 October 2013; Sage Publications Sage CA: Los Angeles, CA, USA, 2013. [[CrossRef](#)]
16. Monk, A. Mode errors: A user-centred analysis and some preventative measures using keying-contingent sound. *J. Man.-Mach. Stud.* **1986**, *24*, 313–327. [[CrossRef](#)]
17. Kolbig, M.; Müller, S. Mode Awareness im Fahrkontext: Eine theoretische Betrachtung. In Proceedings of the 10. Berlinger Werkstatt Mensch-Maschine Systeme, Berlin, Germany, 10–12 October 2013; Brandenburg, E., Doria, L., Gross, A., Günzler, T., Smieszek, H., Eds.; Universitätsverlag der TU Berlin: Berlin, Germany, 2013.
18. Othersen, I. *Vom Fahrer zum Denker und Teilzeitlenker*; Springer: Berlin/Heidelberg, Germany, 2016. [[CrossRef](#)]

19. Brederke, J.; Lankenau, A. Safety-relevant mode confusions—Modelling and reducing them. *Reliab. Eng. Syst. Safe.* **2005**, *88*, 229–245. [[CrossRef](#)]
20. Brederke, J.; Lankenau, A. A rigorous view of mode confusion. In Proceedings of the International Conference on Computer Safety, Reliability, and Security, Catania, Italy, 10–13 September 2002; Anderson, S., Felici, M., Bologna, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2002. [[CrossRef](#)]
21. Sarter, N. Investigating mode errors on automated flight decks: Illustrating the problem-driven, cumulative, and interdisciplinary nature of human factors research. *Hum. Factors* **2008**, *50*, 506–510. [[CrossRef](#)] [[PubMed](#)]
22. Dekker, S. *The field Guide to Understanding Human Error*, 3rd ed.; Ashgate Publishing Ltd.: Aldershot, UK, 2006.
23. Lindberg, T. *Entwicklung Einer ABK-Metapher für Gruppierte Fahrerassistenzsysteme*; Dr. Hut: Berlin, Germany, 2012; pp. 1–329.
24. Rettenmaier, M. Mode Awareness im Fahrzeug: Absicherung von hochautomatisiertem Fahren. Master's Thesis, Technical University Munich, Munich, Germany, 2017.
25. Lee, J.D.; See, K.A. Trust in automation: Designing for appropriate reliance. *Hum. Factors* **2004**, *46*, 50–80. [[CrossRef](#)] [[PubMed](#)]
26. Parasuraman, R.; Riley, V. Humans and automation: Use, misuse, disuse, abuse. *Hum. Factors* **1997**, *39*, 230–253. [[CrossRef](#)]
27. Martinez von Bülow, R.; Raisch, F. Driver assistance systems: Highly automated driving—Acceptance and benefit for customers. In Proceedings of the VDA Technical Congress, Berlin, Germany, 27–28 February 2018.
28. Van Loon, R.J.; Martens, M.H. Automated driving and its effect on the safety ecosystem: How do compatibility issues affect the transition period? *Procedia Manuf.* **2015**, *3*, 3280–3285. [[CrossRef](#)]
29. Hergeth, S.; Lorenz, L.; Vilimek, R.; Krems, J.F. Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. *Hum. Factors* **2016**, *58*, 509–519. [[CrossRef](#)]
30. Parasuraman, R.; Manzey, D.H. Complacency and bias in human use of automation: An attentional integration. *Hum. Factors* **2010**, *52*, 381–410. [[CrossRef](#)]
31. Moray, N.; Inagaki, T. Attention and complacency. *Theor. Issues Ergon.* **2000**, *1*, 354–365. [[CrossRef](#)]
32. Manzey, D.; Reichenbach, J.; Onnasch, L. Human performance consequences of automated decision aids: The impact of degree of automation and system experience. *J. Cogn. Eng. Decis. Mak.* **2012**, *6*, 57–87. [[CrossRef](#)]
33. Forster, Y.; Naujoks, F.; Neukum, A. Your turn or my turn? Design of a human-machine interface for conditional automation. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ann Arbor, MI, USA, 24–26 October 2016. [[CrossRef](#)]
34. Naujoks, F.; Purucker, C.; Wiedemann, K.; Neukum, A.; Wolter, S.; Steiger, R. Driving performance at lateral system limits during partially automated driving. *Accid. Anal. Prev.* **2017**, *108*, 147–162. [[CrossRef](#)]
35. Kromann, C.B.; Jensen, M.L.; Ringsted, C. The effect of testing on skills learning. *Med. Educ.* **2009**, *43*, 21–27. [[CrossRef](#)] [[PubMed](#)]
36. Banks, V.A.; Stanton, N.A. Discovering driver-vehicle coordination problems in future automated control systems: Evidence from verbal commentaries. *Procedia Manuf.* **2015**, *3*, 2497–2504. [[CrossRef](#)]
37. Endsley, M.R. Autonomous driving systems: A preliminary naturalistic study of the Tesla Model S. *J. Cogn. Eng. Decis. Mak.* **2017**, *11*, 225–238. [[CrossRef](#)]
38. Klauer, S.G.; Dingus, T.A.; Neale, V.L.; Sudweeks, J.D.; Ramsey, D.J. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. *NHTSA* **2006**, 594. [[CrossRef](#)]
39. Naujoks, F.; Befelein, D.; Wiedemann, K.; Neukum, A. A review of non-driving-related tasks used in studies on automated driving. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Prague, Czech Republic, 26–28 October 2016; Naujoks, F., Befelein, D., Wiedemann, K., Neukum, A., Eds.; Springer International Publishing: Basel, Switzerland, 2017. [[CrossRef](#)]
40. Feldhütter, A.; Gold, C.; Schneider, S.; Bengler, K. How the duration of automated driving influences take-over performance and gaze behavior. In *Advances in Ergonomic Design of Systems, Products and Processes*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 309–318. [[CrossRef](#)]
41. De Winter, J.C.; Happee, R.; Martens, M.H.; Stanton, N.A. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transp. Res. Part A Traff. Psych. Beh.* **2014**, *27*, 196–217. [[CrossRef](#)]

42. Toffetti, A.; Wilschut, E.S.; Martens, M.H.; Schieben, A.; Rambaldini, A.; Merat, N.; Flemisch, F. CityMobil: Human factor issues regarding highly automated vehicles on eLane. *Transp. Res. Rec.* **2009**, *2110*, 1–8. [[CrossRef](#)]
43. Victor, T.W.; Tivesten, E.; Gustavsson, P.; Johansson, J.; Sangberg, F.; Ljung Aust, M. Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel. *Hum. Factors* **2018**, *60*, 1095–1116. [[CrossRef](#)]
44. Wang, J.; Söffker, D. Bridging gaps among human, assisted, and automated driving with DVIs: A conceptual experimental study. *IEEE T. Intell. Transp. Syst.* **2018**, *20*, 2096–2108. [[CrossRef](#)]
45. Wickens, C.D. Attention to safety and the psychology of surprise. In Proceedings of the 2001 Symposium on Aviation Psychology, Columbus, OH, USA, 5–8 March 2001.
46. Körber, M.; Cingel, A.; Zimmermann, M.; Bengler, K. Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manuf.* **2015**, *3*, 2403–2409. [[CrossRef](#)]
47. Feldhütter, A.; Härtwig, N.; Kurpiers, C.; Hernandez, J.M.; Bengler, K. Effect on mode awareness when changing from conditionally to partially automated driving. In Proceedings of the Congress of the International Ergonomics Association, Florence, Italy, 26–30 August 2019; Springer: Berlin/Heidelberg, Germany, 2019. [[CrossRef](#)]
48. Crundall, D.E.; Underwood, G. Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics* **1998**, *41*, 448–458. [[CrossRef](#)]
49. Land, M.F.; Lee, D.N. Where we look when we steer. *Nature* **1994**, *369*, 742–744. [[CrossRef](#)] [[PubMed](#)]
50. Recarte, M.A.; Nunes, L.M. Effects of verbal and spatial-imagery tasks on eye fixations while driving. *J. Exp. Psychol. Appl.* **2000**, *6*, 31–43. [[CrossRef](#)] [[PubMed](#)]
51. Summala, H.; Nieminen, T.; Punto, M. Maintaining lane position with peripheral vision during in-vehicle tasks. *Hum. Factors* **1996**, *38*, 442–451. [[CrossRef](#)]
52. Jian, J.-Y.; Bisantz, A.M.; Drury, C.G. Foundations for an empirically determined scale of trust in automated systems. *Int. J. Cogn. Ergon.* **2000**, *4*, 53–71. [[CrossRef](#)]
53. Madsen, M.; Gregor, S. Measuring human-computer trust. In Proceedings of the 11th Australasian Conference on Information Systems, Brisbane, Australia, 6–8 December 2000.
54. Åkerstedt, T.; Gillberg, M. Subjective and objective sleepiness in the active individual. *Int. J. Neurosci.* **1990**, *52*, 29–37. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).