Towards standardized metrics for measuring takeover performance in conditionally automated driving: A systematic review

Yining Cao, Feng Zhou, Elizabeth M. Pulver, Lisa J. Molnar, Lionel P. Robert, Dawn M. Tilbury, X. Jessie Yang

University of Michigan[1], State Farm Mutual Automobile Insurance Company[2]

A particular concern with SAE Level 3 automated vehicles is the takeover transition from the automated vehicle to the driver. Prior research has employed a wide range of metrics for measuring takeover performance. However, the lack of a set of standard metrics for measuring takeover performance makes it difficult to consolidate findings and summarize the influence of different factors. This article presents a review of the metrics employed in empirical literature examining takeover transitions in Level 3 automated driving and proposes a framework for standardizing the objective takeover performance metrics.

INTRODUCTION

A particular concern with SAE (Society of Automotive Engineers) Level 3 automation is the takeover transition from the automated vehicle to the driver. As the driver is not always required to monitor the environment, s/he may become increasingly decoupled from the driving task and have difficulty taking over control at a moment's notice (Ayoub, Zhou, Bao, & Yang, 2019; Du et al., 2020; Zhou, Yang, & Zhang, 2020).

To measure takeover performance, researchers have identified and employed a wide range of metrics in both time and quality aspects. However, the lack of standard metrics for measuring takeover performance makes it difficult to consolidate findings from prior studies, summarize the influence of different factors on takeover performance, and compare various designs aimed to facilitate takeover transitions (McDonald et al., 2019; Zhang et al. 2019).

This paper aims to propose a framework for standardizing the objective takeover performance metrics. A review of objective metrics employed in empirical literature examining takeover transitions in Level 3 automated driving is presented.

METHOD

Databases and Search strategy

We performed a thorough search of related studies following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) process (Moher et al., 2009) across different databases, including Association for Computer Machinery (ACM) digital library, IEEE Xplore, Scopus, ScienceDirect, PsycINFO, PsycArticle and Transportation Research International Document (TRID). The two sets of domain-specific searching key terms are:

Set1 = {"automated vehicle", "automated driving", "autonomous vehicle}

Set2 = {"takeover", "handover", "control transition"}

We used a combination of two key terms (one in each set) for searching. Key terms were searched within metadata (i.e., title, abstract, and keywords). Papers were restricted to journal publications, conference proceedings, and theses. Eligible studies were within a period from January 2009 to January 2020.

In total, 1378 papers were identified in the selected database and 15 additional papers were identified by scanning references in review papers and asking fellow researchers for relevant studies. We removed duplicates and identified 595 unique papers in total.

Eligibility criteria

All identified papers were screened for eligibility based on the following three criteria:

1. The paper reported takeover performance in a human-subjects experiment.

2. In the experiment, participants had to take over control from Level 3 driving automation.

3. At least one takeover quality aspect metric and one-time aspect metric were reported.
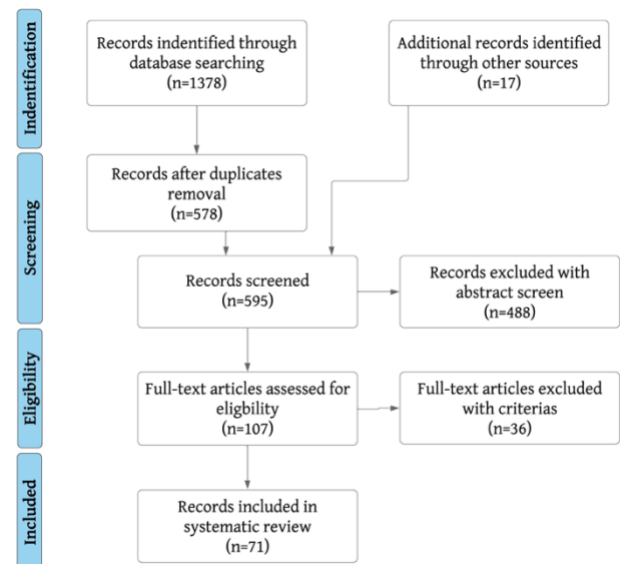
Study selection and data extraction

Figure 1. PRISMA overview method for literature searching

After scanning all 595 papers for eligibility based on abstract and methodology, we considered 107 relevant papers for further assessment in accordance with the criteria above. In the full-text article assessment stage, 36 papers that did not strictly follow the criterion were excluded for the following reasons: no time aspect metric was measured (7 papers), no objective quality aspect metric was measured (21 papers), not

a Level 3 driving automation system (3 papers), content duplicate (i.e., a conference paper and a journal article describing the same study, 5 papers). In the end, we included 71 papers in the systematic review. The paper selection process is shown in Figure 1.

RESULTS

A Framework for Quantifying Takeover Performance

We proposed a framework (Figure 3) that contains objective metrics for quantifying takeover performance and conducted a large-scale literature review to synthesize existing studies. The framework contains both time and quality aspects metrics. All metrics reviewed can be mapped to the framework.

For the time aspects metrics, we propose and define a three stages takeover process model based on the takeover time sequence (Figure 2). The takeover process begins at the moment of takeover request (TOR) issued. In the orientation stage, Drivers perform intuitive actions without full awareness of the takeover situation. In this stage, drivers notice the TOR and switch attention from the Non-driving related task (NDRT) to the takeover situation. After they gained enough information and decided on action selection, the initialization stage starts. This stage is for measuring the timeliness of the driver's conscious actions and information assimilation. The following action execution stage measures driving performance after the maneuver initialized. This stage starts from a predefined point of maneuver initiation to the end of maneuver execution.
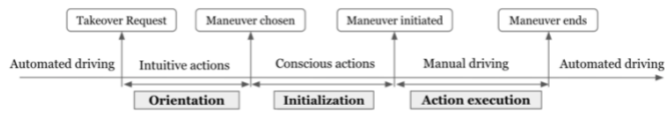

Figure 2. Three stages of the takeover process

Takeover quality consists of various components related to takeover maneuvers such as speed, wheel, and pedal behavior (Wu et al., 2019) and also margins between the ego vehicle and potential obstacles indicating the takeover safeness. (McDonald et al., 2019).
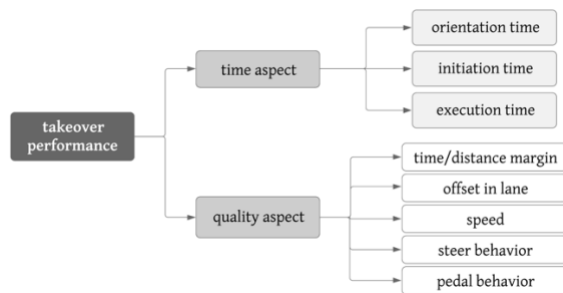

Figure 3. The proposed framework for takeover performance metrics

Time Aspect Metrics

The empirical literature has reported various takeover time aspect metrics with different starting and ending points. According to our framework, drivers will go through orientation, initiation, and action execution process to regain

vehicle control after TOR. Time aspect metrics are reviewed according to these three stages (See Table 1).

Table 1. Takeover time aspect metrics: Measurement, description, and number of papers employing the corresponding measurement

Measurement	Description	# of Papers
Orientation Stage		
Gaze reaction time	Time interval between TOR and the first saccade stirs from the instrument cluster	3
Road fixation time	Time interval between TOR and the first fixation at the scenery	6
Forward-gaze time	Time interval between TOR and initial indication of when driver first gazes in a forward direction	1
Movement time	Time interval between TOR and the driver started moving left or right	2
Initialization Stage		
Side-mirror time	Time interval between TOR and the driver's first glance at the side mirror	3
Eye-on-windshield reaction time	Time interval between TOR and first gaze on the windshield	1
Hands-on time/Foot-on-pedal time	Time interval between TOR and hand on the steering wheel or foot returned to the pedal	14
Notification response time	Time between TOR and switch off of automation	30
Takeover time	Time interval between TOR and the driver's first maneuver	25
Action Execution Stage		
Lane changing time	Time interval between TOR and lane changing maneuver finished	3

Orientation Stage. In the orientation stage, drivers begin intuitive responses to takeover situations without maneuver decisions made. Metrics in this stage are used to measure how fast drivers can switch attention, assimilate information, and gain readiness. Four metrics have been reported in the orientation stage. Gaze reaction time, road fixation time, and forward-gaze time are used to measure cognitive readiness in the orientation stage. Gaze reaction time is the time until the first saccade stirs away from the Non-driving related task (NDRT) (Gold et al., 2013). This is the first indication of the driver's awareness of the takeover situation. Road fixation time is defined as the time interval from TOR to the first fixation on the road. This measurement is important as it indicates how fast the drivers begin paying attention to the road condition and being ready to drive (Yoon & Ji, 2019). Abe, Sato, Uchida, and Itoh (2019) also mentioned a similar metrics called forward-gaze time, which is the time between the initial indication to the first time when the driver gazes in a forward direction. Movement time is used to indicate a driver's motor readiness. Movement time is defined as the time until the driver starts moving left or right hand, used as the first indication of drivers' physical readiness (Gold et al., 2013; Kerschbaum et al., 2015). This metric is usually used in driving scenarios where the NDRT is dynamically operated by hand.

Initialization Stage. Metrics in the initialization stage is related to conscious intervention after drivers arrive at a decision and start the maneuver, usually indicated by the performance of first conscious input of either braking or

steering (Gold et al., 2013; van der Heiden, Iqbal, & Janssen, 2017). For measuring the cognitive performance in this stage, gaze reactions are used as indicators of drivers' situational awareness. Side mirror time is the first gaze to the left side mirror, necessary to ensure a safe lane change (Gold et al., 2013). Eriksson et al. (2018) mentioned the eyes on windshield reaction time as a similar measurement to side mirror time. This metric is defined as the time between TOR and the moment when the eye gaze of the driver first is detected in the windshield area. A shorter side mirror time can be an indication of a faster information assimilation process. Hands-on time, which is also named as preparation for action time (Vogelpohl et al., 2018), measures the fastness of the driver has his/her hand on the steering wheel (or foot returned to the pedal in the braking situation). The hands-on time represents the physical readiness of the driver to reengage control (Yoon et al., 2019). Takeover time, the most widely used time metric, measures the time interval between the TOR and a predefined point indicating initiation of conscious maneuver. This metric helps understand a driver's ability to react (Gold, Happee, & Bengler, 2018). The thresholds of predefined points for brake pedal input and steer angle vary across different studies. The most frequently used are 2° change of steering wheel angle and 10% actuation of the brake pedal. However, some studies do not explicitly define the threshold of the initialization points. They used the time interval between TOR and the switch-off of the automation system by activating the actuators (buttons, pedal, wheel, etc.) to indicate initialization. We differentiate it from the takeover time metric as notification response time.

Action Execution Stage. Time metrics in this stage indicate the total time used to execute the entire maneuver, such as the time for completing the first lane change (Telpaz, Rhindress, Zelman, & Tsimhoni, 2015). Time is calculated between the moment of TOR and the point indicating the end of the maneuver (e.g., the center of ego vehicle crosses the line boundary; Dogan et al., 2019; Eriksson et al., 2018).

Quality Aspect Metrics

Takeover quality can be assessed by evasive maneuver and stabilization (Körber et al., 2018). We categorized metrics measuring takeover quality into five different aspects: time/distance margin, speed, offset in lane, steering and brake behavior. Related metrics are summarized in Table 2.

Table 2. Takeover quality aspect metrics: Measurement, description, and number of papers employing the corresponding measurement

Measurement	Description	Statistics (# of Papers)
Time/Distance margin		
Time to Collision (TTC)	Time budget to potential collision to the obstacle ahead. Usually measured with defined points or minimal value	min (17); abs*(3)
Time to Lane Crossing (TLC)	Time budget to potential collision to lane boundary	min (2)
Distance to Collision	Distance to the obstacle	min (5)
Time Head Way (THW)	Time interval between two vehicles	min (3)
Number of Collision	Number of collision accident happened	count(7)
Speed		
Longitudinal Speed	Velocity statistics calculated in the manual driving period after takeover	mean (11); max (7); sd. (7); min (1)
Longitudinal Acceleration	Acceleration/deceleration statistics calculated in the manual driving period after takeover. Resulting acceleration is the geometric mean of the longitudinal and latitudinal acceleration	max (7); mean (6); min(1); sd.(1)
Latitudinal Acceleration		max(5); avg.(2)
Resulting Acceleration		max(3)
Resulting Jerk	Jerk is derivative of acceleration. Resulting jerk is the geometric mean of the longitudinal and latitudinal jerk	max(1)
Offset in lane		
Offset from lane center	Statistics measuring the deviation of the ego-vehicle from the center of the ego-lane during the manual driving period	sd.(19); max(7); mean (2)
Steering behavior		
Steering angle velocity	Statistics measuring the steering angle velocity during the takeover operation	max(3); sd.(2)
Steering Angle Reversal	Steering reversals from the start of the event to the end of the event	rate/count(8)
Steering position	Statistics measuring the steering wheel angle during manual driving period	sd.(11); max(8); entropy(1)
Steering Angle Acceleration	The standard deviation of steering angular acceleration	sd.(1)
Brake behavior		
Brake pedal input	Statistics measuring the brake pedal input during the takeover operation	count/rate(3); max(1)
Brake pedal speed	The maximum pedal speed during takeover	max(1)

* abs: absolute value calculated with defined points.

Time/Distance margin. These metrics are used to measure the possibility of collision in avoidance scenarios. Among all the collision metrics, TTC is the most commonly used. It measures the remaining time until the evasive maneuver ends (i.e., full stop or change to another lane), which can alternatively be termed as 'remaining action time' (Gold et al., 2013). The calculation assumes constant speed of both the ego-vehicle and the obstacle (Gold et al., 2018). The Minimum TTC was calculated within a certain time sequence, usually between the TOR and the point when the evasive maneuver ended. Longer minimum TTC is an indication of a safer takeover (Borojeni, Weber, Heuten, & Boll, 2018; Dogan et al., 2017; Hergeth et al., 2017; Zhang, de Winter, et al., 2019). Most studies used minimal TTC, while some used TTC directly by defining the start and end points. For example, Borojeni et al. (2018) defined the calculation point as the first action (i.e., steering or braking). In scenarios without imminent collision, time to lane crossing (TLC) rather than TTC was used (Braunagel et al., 2017; Zeeb et al., 2016). TLC was defined as the time budget to a potential collision with a lane boundary. THW is another measurement for potential collision. This metric is calculated between two vehicles in sequence as regulated in SAE J2944 (SAE, 2015). Besides metrics related to time budget, some studies also used

distance budget (Braunagel et al., 2017) or the number of collisions (Dogan et al., 2019) to measure safety of takeover.

Speed. Speed metrics measurement consists of vehicle velocity and acceleration. Specific statistics are calculated during the manual driving interval. Acceleration reflects the forces the tires had to transfer and the performance of immediacy (Kim & Yang, 2017). If the acceleration approaches the physical limit, the driving condition becomes unstable. Thus, the maximum acceleration that occurs after TOR is considered a good measure for the quality of reaction (Gold et al., 2013). Jerk is a measurement used to evaluate shift quality and ride comfort, with a smaller value representing higher takeover quality (Du et al., 2020).

Offset in lane. Lane offset is a strong indicator of lateral control. The standard deviation of lane position (SDLP) was the most commonly reported metric which described the dispersion of the lateral lane position. Both the mean and maximum lane departure distance measured relative to the center of the lane were used as indicators of the lane-keeping intention (Kim & Yang, 2017).

Steering behavior. Steering behaviors describe the smoothness of the maneuver (Lindemann et al., 2019). Steering reversal is defined as a situation in which the steering wheel rotates at least a specified amount in one direction and then rotates at least an equal amount in the opposite direction within a time window. A higher rate of steering reversal indicates an increasing driver distraction window (SAE, 2015). In calculating the steer angle reversal rate, the angular rotation threshold Δa and the time window Δt should be defined. Steering entropy is a dimensionless value between 0 and 1. It measures the consistency/randomness of the steering wheel angle as an indicator of a driver's workload. Higher entropy is likely due to driver distraction (SAE, 2015). Kamezaki et al. (2019) compared the steering entropy for manual driving and manual takeover and concluded that manual takeover led to a sudden surge of driver workload.

Brake behavior. Pedal behavior indicates the quickness and intensity of takeover maneuvers. Sudden and intense braking is considered dangerous in naturalistic driving situations. Thus, some studies used the maximum value of the brake pedal input as an indicator of takeover performance. Roche and Brandenburg (2018) found that maximum brake pedal input was significantly higher with a smaller takeover time budget and higher TOR urgency.

DISCUSSION AND CONCLUSION

The review represents a thorough examination of metrics that have been used in previous takeover-related studies. We found a lack of consistency in the naming conventions for various metrics and their calculation method, which makes it difficult for direct comparisons among studies.

Inconsistent metrics calculation. Metrics calculation needs to be rigorously defined with a clear specification of parameters and time windows. In previous studies, the calculations were inconsistent. For example, when calculating takeover time, the most frequently used metric for measuring takeover timeliness, the starting point was consistently defined as the time of TOR, while the ending point varied as shown in Table

3. Differences were also apparent in definitions of steering reversal rate, TTC, SDLP, etc.

Table 3. Different Definitions of Endpoints for Takeover Time Calculation

End point for takeover time calculation	Reference
2° steering wheel angle/10% braking pedal position	Gold et al., 2013
10° steering wheel angle	Zeeb et al., 2016
0.036% braking pedal position	Hergeth et al., 2017
0.25° steering wheel angle	Petermeijer et al., 2017
8° steering wheel angle/0.1% braking pedal position	Roche & Brandenburg et al., 2018
5° steering wheel angle	Kamezaki et al., 2019
3° steering wheel angle	Kunze et al., 2019
1° steering wheel angle/1% braking pedal position	Lindemann et al., 2019
1° steering wheel angle/10% braking pedal position	Wu et al., 2019
1.4° steering wheel angle/9.2% braking pedal position/3.2% throttle pedal	Roche et al., 2019

Inconsistent naming conventions. The same metrics were given different names across studies. For example, the term "reaction time" was applied to different processes. Kim and Yang (2017) referred it to the time interval between TOR and the initiation of maneuver, while Schmidt et al. (2017) defined it as hands-on time.

In this article, we present a consistent naming convention and discussed the definitions of metrics according to the identified papers. Further research is needed to develop standard documentation for regularizing the measurements of takeover performance metrics.

ACKNOWLEDGMENT

REFERENCES

Abe, G., Sato, K., Uchida, N., & Itoh, M. (2019). Effect of Changes in Levels of Automated Driving on Manual Control Recovery. *IFAC-PapersOnLine, 52(19)*, 79-84.

Ayoub, J., Zhou, F., Bao, S., & Yang, X. J. (2019, September). From manual driving to automated driving: A review of 10 years of autoui. In *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications* (pp. 70-90).

Borojeni, S. S., Weber, L., Heuten, W., & Boll, S. (2018, September). From reading to driving: priming mobile users for take-over situations in highly automated driving. In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services* (pp. 1-12).

Dogan, E., Honnêt, V., Masfrand, S., & Guillaume, A. (2019). Effects of non-driving-related tasks on takeover performance in different takeover situations in conditionally automated driving. *Transportation research part F: traffic psychology and behaviour*, 62, 494-504.

Dogan, E., Rahal, M. C., Deborne, R., Delhomme, P., Kemeny, A., & Perrin, J. (2017). Transition of control in a partially automated vehicle: Effects *of anticipation and non-driving-related task*

involvement. Transportation research part F: traffic psychology and behaviour, 46, 205-215.

Du, N., Zhou, F., Pulver, E. M., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Predicting driver takeover performance in conditionally automated driving. *Accident Analysis & Prevention, 148*, 105748.

Eriksson, A., Petermeijer, S. M., *Zimmermann, M., De Winter, J. C., Bengler, K. J., & Stanton, N. A. (2018). Rolling out the red (and green) carpet: supporting driver decision making in automation-to-manual transitions. IEEE Transactions on Human-Machine Systems, 49(1)*, 20-31.

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013, September). "Take over!" How long does it take to get the driver back into the loop? In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 57, No. 1, pp. 1938-1942).

Gold, C., Happee, R., & Bengler, K. (2018). Modeling take-over performance in level 3 conditionally automated vehicles. *Accident Analysis & Prevention, 116*, 3-13.

Gold, C., Körber, M., Lechner, D., & Bengler, K. (2016). Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density. *Human factors*, *58*(4), 642-652.

Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. *Human factors*, *59*(3), 457-470.

Kamezaki, M., Hayashi, H., Manawadu, U. E., & Sugano, S. (2019). Human-centered intervention based on tactical-level input in unscheduled takeover scenarios for highly-automated vehicles. *International Journal of Intelligent Transportation Systems Research*, 1-10.

Kim, H. J., & Yang, J. H. (2017). Takeover requests in simulated partially autonomous vehicles considering human factors. *IEEE Transactions on Human-Machine Systems*, *47*(5), 735-740.

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied ergonomics*, *66*, 18-31.

Kunze, A., Summerskill, S. J., Marshall, R., & Filtness, A. J. (2019, September). Conveying uncertainties using peripheral awareness displays in the context of automated driving. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 329-341).

Lindemann, P., Müller, N., & Rigolll, G. (2019, June). Exploring the Use of Augmented Reality Interfaces for Driver Assistance in Short-Notice Takeovers. In *2019 IEEE Intelligent Vehicles Symposium (IV)* (pp. 804-809). IEEE.

McDonald, A. D., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., & Yuma, N. (2019). Toward computational simulations of behavior during automated driving takeovers: a review of the empirical and modeling literatures. *Human factors*, *61*(4), 642-688.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097.

Mok, B. K. J., Johns, M., Lee, K. J., Ive, H. P., Miller, D., & Ju, W. (2015, June). Timing of unstructured transitions of control in automated driving. In *2015 IEEE intelligent vehicles symposium (IV)* (pp. 1167-1172). IEEE.

Society of Automotive Engineers. (2015). Operational definitions of driving performance measures and statistics.

Petermeijer, S., Doubek, F., & de Winter, J. (2017, October). Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1505-1510). IEEE.

Roche, F., & Brandenburg, S. (2018, November). Should the urgency of auditory-tactile takeover requests match the criticality of takeover situations? In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1035-1040). IEEE.

Roche, F., Somieski, A., & Brandenburg, S. (2019). Behavioral changes to repeated takeovers in highly automated driving: effects of the takeover-request design and the nondriving-related task modality. *Human factors*, *61*(5), 839-849.

Schartmüller, C., Riener, A., Wintersberger, P., & Frison, A. K. (2018, September). Workaholistic: on balancing typing-and handover-performance in automated driving. In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services* (pp. 1-12).

Schmidt, J., Dreißig, M., Stolzmann, W., & Rötting, M. (2017, September). The influence of prolonged conditionally automated driving on the take-over ability of the driver. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1974-1978). Sage CA: Los Angeles, CA: SAGE Publications.

Telpaz, A., Rhindress, B., Zelman, I., & Tsimhoni, O. (2015, September). Haptic seat for automated driving: preparing the driver to take control effectively. In *Proceedings of the 7th international conference on automotive user interfaces and interactive vehicular applications* (pp. 23-30).

van der Heiden, R. M., Iqbal, S. T., & Janssen, C. P. (2017, May). Priming drivers before handover in semi-autonomous cars. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 392-404).

Vogelpohl, T., Kühn, M., Hummel, T., Gehlert, T., & Vollrath, M. (2018). Transitioning to manual driving requires additional time after automation deactivation. *Transportation research part F: traffic psychology and behaviour, 55*, 464-482.

Wu, Y., Kihara, K., Takeda, Y., Sato, T., Akamatsu, M., & Kitazaki, S. (2019). Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving. *Accident Analysis & Prevention, 124*, 202-209.

Yoon, S. H., & Ji, Y. G. (2019). Non-driving-related tasks, workload, and takeover performance in highly automated driving contexts. *Transportation research part F: traffic psychology and behaviour, 60,* 620-631.

Yoon, S. H., Kim, Y. W., & Ji, Y. G. (2019). The effects of takeover request modalities on highly automated car control transitions. *Accident Analysis & Prevention, 123,* 150-158.

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident analysis & prevention, 78,* 212-221.

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident analysis & prevention, 92,* 230-239.

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation research part F: traffic psychology and behaviour*, *64*, 285-307.

Zhou, F., Yang, X. J., & Zhang, X. (2020). Takeover transition in autonomous vehicles: a YouTube study. *International Journal of Human–Computer Interaction, 36(3),* 295-306.