Determinants of take-over time from automated driving
A meta-analysis of 129 studies

Zhang, Bo; de Winter, Joost; Varotto, Silvia; Happee, Riender; Martens, Marieke

DOI
10.1016/j.trf.2019.04.020

Publication date
2019

Document Version
Final published version

Published in
Transportation Research Part F: Traffic Psychology and Behaviour

Citation (APA)
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. https://doi.org/10.1016/j.trf.2019.04.020

Transportation Research Part F 64 (2019) 285–307

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Determinants of take-over time from automated driving: A meta-analysis of 129 studies

Bo Zhang [a,*], Joost de Winter [b], Silvia Varotto [c,d], Riender Happee [b,c], Marieke Martens [a,e]

[a] Centre for Transport Studies, University of Twente, Enschede, the Netherlands
[b] Department of Cognitive Robotics, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands
[c] Transport & Planning, Faculty of Civil Engineering and Geoscience, Delft University of Technology, Delft, the Netherlands
[d] SWOV Institute for Road Safety Research, The Hague, the Netherlands
[e] TNO Traffic & Transport, The Hague, the Netherlands

A R T I C L E I N F O

Article history:
Received 18 May 2018
Received in revised form 13 March 2019
Accepted 16 April 2019
Available online 31 May 2019

Keywords:
Automated driving
Transition of control
Take-over time
Literature review

A B S T R A C T

An important question in automated driving research is how quickly drivers take over control of the vehicle in response to a critical event or a take-over request. Although a large number of studies have been performed, results vary strongly. In this study, we investigated mean take-over times from 129 studies with SAE level 2 automation or higher. We used three complementary approaches: (1) a within-study analysis, in which differences in mean take-over time were assessed for pairs of experimental conditions, (2) a between-study analysis, in which correlations between experimental conditions and mean take-over times were assessed, and (3) a linear mixed-effects model combining between-study and within-study effects. The three methods showed that a shorter mean take-over time is associated with a higher urgency of the situation, not using a handheld device, not performing a visual non-driving task, having experienced another take-over scenario before in the experiment, and receiving an auditory or vibrotactile take-over request as compared to a visual-only or no take-over request. A consistent effect of age was not observed. We also found the mean and standard deviation of the take-over time were highly correlated, indicating that the mean is predictive of variability. Our findings point to directions for new research, in particular concerning the distinction between drivers' ability and motivation to take over, and the roles of urgency and prior experience.

1. Introduction

Until automated driving systems are capable of performing all driving tasks under all road conditions (i.e., full automation as defined by SAE International, 2016), drivers will have to take over control when the automation fails or reaches its operational limits. Partial automation (SAE L2), which is already made available by several car manufacturers, requires drivers to monitor the road and to be prepared for immediate intervention in case of critical events. At higher levels of automation (SAE L3 'conditional automation' and L4 'high automation'), drivers are allowed to engage in non-driving activities, while the automation executes the monitoring task and issues a take-over request (TOR) when the driver has to intervene. How long it takes drivers to reclaim manual control and what factors determine take-over time are important questions for both scientific researchers and automobile manufacturers.

* Corresponding author at: Drienerlolaan 5, 7522 NB Enschede, the Netherlands.
E-mail address: b.zhang@utwente.nl (B. Zhang).

https://doi.org/10.1016/j.trf.2019.04.020
1369-8478/

1.1. Driver take-over process and response times

The driver take-over process comprises several information-processing stages: perception of visual, auditory, and/or vibrotactile stimuli, cognitive processing of the information, response selection (decision making), resuming motor readiness (by repositioning the hands and feet on the steering wheel and pedals), and the actual action (e.g., steering and braking input to the vehicle) (Gold & Bengler, 2014; Gold, Damböck, Lorenz, & Bengler, 2013; Petermeijer, De Winter, & Bengler, 2016; Zeeb, Buchner, & Schrauf, 2015; Zhang, Wilschut, Willemsen, & Martens, in press). Gold, Damböck et al. (2013) described four response (RT) measures: (1) gaze response time, (2) eyes-on-road time, (3) hands-on-wheel response time, and (4) take-over time (i.e., intervention time). In addition, researchers have used task-specific measures, such as hand-movement response time (e.g., Kerschbaum, Lorenz, & Bengler, 2015; Kerschbaum, Omozik, Wagner, Levin, Hermsdörfer, & Bengler, 2017; Zhang et al., in press), mirror check response time (e.g., Gold, Damböck et al., 2013; Vogelpohl, Kühn, Hummel, Gehlert, & Vollrath, 2018) and lane change response time (e.g., Petermeijer, Cieler, & De Winter, 2017; Eriksson et al., 2019). Although different response time measures can be distinguished, take-over time (TOT), defined as the time that drivers take to resume control from automated driving after a critical event in the environment or after having received a TOR, appears to be the most frequently used measure in the literature.

The temporal sequence of the take-over process is illustrated in Fig. 1. Typically, the driver has to take over within the 'time budget' available until the system limit of the automation is reached. Such system limits may comprise an upcoming collision (e.g., with a stationary vehicle in the ego lane) or operational limits of the automated driving system (e.g., due to missing lane markings). If drivers do not take over within the available time budget, serious safety issues may occur.

1.2. Previous review studies

The empirical literature reports a wide range of TOT values. For example, De Winter, Stanton, Price, and Mistry (2016) reported a mean TOT of 0.87 s (*SD* = 0.24 s) when the participants were required to brake in response to a salient red stop sign, whereas, in Politis et al. (2018), participants took over control on average 19.8 s (*SD* = 9.3 s) after the onset of a 60 s countdown TOR.

Several researchers have provided narrative reviews of TOT studies. Radlmayr and Bengler (2015) summarised eleven studies that investigated the effect of take-over time budget and concluded that longer time budgets are associated with longer TOTs and better take-over quality. A time budget smaller than 7 s was regarded as insufficient for a fully distracted driver to successfully take over control. Vogelpohl, Vollrath, Kühn, Hummel, and Gehlert (2016) provided an overview of 22 TOT studies involving a transition from highly automated driving (SAE L3) to manual driving and identified potentially influential factors related to the environment, the driver, the human-machine interface, and the vehicle. The authors suggested that the complexity of the take-over situation, the modality of the TOR, and the non-driving task (NDT) performed at the moment of the TOR are important factors. Furthermore, a traffic situation of high complexity and engagement in NDTs were argued to lead to slow responses, whereas multi-modal TORs shortened the TOT and improved the take-over quality. In another literature survey, Walch et al. (2017) discussed 17 take-over studies, focusing on the effect of the time budget, traffic complexity, NDT, and driver age. The authors concluded that 10 s seems an adequate time budget, while pointing out that the driver state and situational circumstances affect the driver's ability to take over control. Vogelpohl et al. (2016) and Walch et al. (2017) both noted that outcomes were sometimes inconsistent between the surveyed studies. For example, it was observed that Gold, Berisha, and Bengler (2015) and Petermann-Stock, Hackenberg, Muhr, and Mergl (2013) reported significantly longer TOTs when the participants were engaged in visual-motor NDTs compared to cognitive-auditory NDTs, whereas this effect was not statistically significant in the experiment by Radlmayr, Gold, Lorenz, Farid, and Bengler (2014). Another example of an inconsistency is that a negative effect of higher traffic complexity on TOT was found in Radlmayr et al. (2014) and Gold, Lorenz, and Bengler (2014), but not in Shen and Neyens (2014). This heterogeneity suggests that a larger number of studies need to be reviewed to draw reliable conclusions.

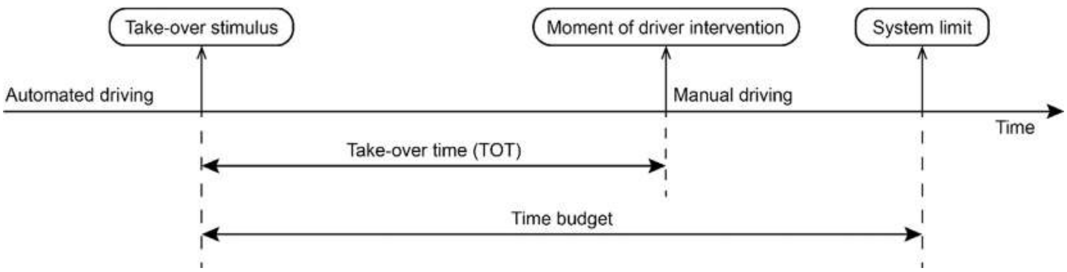

Fig. 1. Illustration of the take-over procedure. The present meta-analysis focuses on the take-over time (TOT), defined as the time between the take-over stimulus (take-over request or critical event in the environment) and the intervention by the driver.

Although a number of narrative reviews exist, little effort has been devoted to quantitatively synthesising the available TOT studies. Eriksson and Stanton (2017) reviewed 25 take-over studies; they extracted 43 take-over time budgets (lead times) which varied between 0 and 30 s (mean = 6.37, *SD* = 5.36 s), and 87 TOTs from 1.14 s to 15 s (mean = 2.96, *SD* = 1.96 s). The authors noted that 3 s, 4 s, 6 s, and 7 s were the most frequently used time budgets and that the corresponding mean TOTs were 1.14, 2.05, 2.69, and 3.04 s. Apart from the time budget, Eriksson and Stanton (2017) did not review the effect of study variables that may affect the TOT. Gold, Happee, and Bengler (2018) provided a predictive model of TOTs based on the datasets obtained from six driving simulator experiments. Out of the seven variables considered in the model, the time budget, traffic density, and repetition (i.e., prior experience) turned out to be significant predictors, whereas driver age, physical and cognitive load of NDT, and the lane in which the ego car was driving (i.e., left, right, or middle) showed only minor effects. A limitation of Gold, Happee et al. (2018) is that only six experiments were analysed and that the experimental settings were similar (i.e., in all the experiments, the take-over scenario was represented by two crashed vehicles blocking the ego lane).

The above reviews suggest that the time budget potentially affects TOTs. However, the existing reviews have several limitations. First, the available reviews analysed only a small number of study variables. Second, most reviews did not numerically synthesise the effects of study variables. Third, the number of reviewed studies is small: The maximum number of studies reviewed was 25 (Eriksson & Stanton, 2017), while this study included 129 TOT studies.

1.3. Research objectives

As pointed out above, there is a need for a new quantitative synthesis of the various TOT studies, having a higher statistical power (i.e., a larger number of included studies) and broader scope (i.e., multiple variables examined simultaneously) as compared to previous reviews. We conducted a comprehensive search of empirical studies and employed meta-analytic methods to examine the predictors of TOT.

Cronbach (1975) discussed two disciplines of scientific psychology: experimental psychology, which is concerned with studying the effects of experimental manipulations, and correlational psychology, which is concerned with understanding differences between individuals and groups. In his work, Cronbach called for combining these two disciplines. A similar approach was followed in the present paper.

First, a within-study meta-analysis was performed to summarise studies that compared pairs of experimental conditions. The within-study analysis describes how mean TOTs are affected by a particular study variable when holding all other study variables constant, thus allowing for statements about causal effects.

Second, because individual experiments typically manipulate only a small number of variables, while experimental conditions differ across studies, we also examined the associations between experimental conditions and TOTs. The second approach concerned a correlation analysis to examine the relationships between the mean TOTs and a comprehensive list of study variables (related to the driver, the automation system, the human–machine interface, the take-over situation, and the experimental set-up) across all studies. The between-study analysis allows for predicting under which experimental conditions the mean TOTs will be low or high.

Third, the within-study experimental approach and the between-study correlational approach were united in a linear mixed-effects model. The mixed model allows for a powerful analysis of the effects of study variables while controlling for the confounding effect of the other study variables.

At the end of the paper, we discuss the similarity in the outcomes of the three methods. Consistent results across all three methods suggest high robustness and generalizability.

2. Methods

This study was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009). The PRISMA checklist is available in the supplementary materials. No protocol was generated or registered.

2.1. Information sources and search strategy

We conducted a literature search with the aim to retrieve as many take-over studies as possible, including grey literature records to minimize publication bias (Rothstein, Sutton, & Borenstein, 2005). Multiple search strategies were used, such as database searching (Google Scholar, ResearchGate), scanning the reference lists of papers, using the 'cited by' feature of Google Scholar, snowballing strategies (Jalali & Wohlin, 2012), and asking fellow researchers for relevant studies. Most searches were conducted using Google Scholar, as it is the most comprehensive search engine, especially for works of the 21st century (De Winter, Zadpoor, & Dodou, 2014; Martín-Martín, Orduna-Malea, Thelwall, & López-Cózar, 2018). Typically used keywords were 'takeover', 'take over', and 'transition of control' in combination with domain-specific keywords, to minimize false positives (e.g., 'automated driving', 'driverless', or the names of often-cited authors such as 'Bengler' or 'Merat'). The searches were performed between October 2, 2016 and December 17, 2018.

2.2. Eligibility criteria

To be included in this review, studies had to fulfil the following six criteria:

1. The study had to involve a transition from partially, conditionally, or highly automated driving (i.e., SAE L2 and above; hands off the steering wheel and feet off the pedals) to manual driving.
2. The study had to involve an automation-to-manual take-over performed by a human (e.g., braking, steering, button pressing).
3. The study had to involve a transition in response to a TOR or a critical event in the environment. That is, this meta-analysis includes only 'mandatory driver-in-control (DC) transitions' as defined by Lu, Happee, Cabrall, Kyriakidis, and De Winter (2016). Studies in which more than one take-over stimulus (i.e., both a TOR and a critical event in the environment) was presented at different moments were not included, because in such cases it cannot be determined to which stimulus the participants responded. For example, in Körber, Baseler, and Bengler (2018), the obstacle was visible three seconds before the TOR, and the drivers were, therefore, able to take over control before the TOR.
4. The study had to report a mean or median take-over time (TOT), or the mean/median TOT should be calculable from the information reported.
5. In the text of the paper, the TOT had to be defined as the time interval between the initiation of the take-over stimulus (i.e., the onset of the TOR or the start of an environmental event that can initiate driver take-over) and the moment of driver intervention (by means of braking, steering, or button pressing).
6. The study had to be written in English or German.

All types of studies were eligible, including journal publications, papers from conference proceedings, theses, reports, posters, and presentation slides. If a publication contained more than one experiment, we considered each experiment as a separate study. We applied no restriction on the year of publication.

2.3. Study selection and data extraction

After initial scanning and filtering, we retrieved 299 potentially relevant full-text records, which were further reviewed for eligibility. After removing 30 duplicate records and 1 record written in a language other than English or German (Japanese), 149 records were excluded for the following reasons: no TOT was measured (103 records), no mean/median TOT was reported (8 records), TOT was not defined according to the fifth eligibility criterion or the TOT measurement was not clearly described (18 records), participants were required to have at least one hand on the wheel during automated driving (4 records), or multiple take-over stimuli were presented at different moments (16 records). In the end, 129 studies from 119 records met the inclusion criteria. These studies comprised 520 mean or median TOT observations. In studies where mean TOT values were only available in figures, the numerical values were extracted using the online tool WebPlotDigitizer (Rohatgi, 2017). Besides the mean and median TOT, which are measures of central tendency and our primary meta-analysis outcome of interest (see Section 1.1), we also extracted the standard deviation of the TOT as an index of variability.

The data extraction and variable coding (described in Section 2.4.2) were conducted by the first author. The second author supervised the process using a file hosting service (Dropbox), conducted multiple manual inspections of the annotated values, and corrected errors.

Ten of the included studies reported no mean TOT, but only the median TOT. Since the distribution of human response times is right-skewed, using unadjusted medians together with means would induce bias. To reduce this bias, we applied a multiplication factor of 1.123 to the median TOTs to obtain an estimate of the corresponding mean TOTs. This correction factor was established from the means and medians from 14 included studies in which both values were reported.

An examination of the included studies showed similar experimental methods. That is, almost all studies used a virtual driving simulator and measured the TOT from the simulator sensor data (i.e., brake pedal depression or steering wheel angle). Although the studies involved simulators of different fidelity levels (e.g., motion base vs. no motion base), different experimental designs (e.g., between-subjects vs. within-subjects), and different experimental protocols (in terms of e.g., participant training, instruction, duration, and breaks), these differences provided no meaningful basis for assessing the study quality. Hence, we did not code the quality of individual studies and considered them of equal importance.

Furthermore, we did not apply weights that depend on sample size (e.g., Hedges & Vevea, 1998; Schmidt & Hunter, 2015). A preliminary analysis showed a substantial skewness of the sample size distribution, with a few large-sample studies including over 100 participants and many moderate-sample studies of about 20 participants. The use of unit weights has been recommended when the sample sizes are unequal, to avoid that the meta-analysis outcome is dominated by a small number of large-sample studies (Osburn & Callender, 1992). Our choice for unit weights is in line with simulation studies showing that unit weights offer similar or sometimes even superior predictive validity as compared to procedures that involve weighting (Bobko, Roth, & Buster, 2007; Einhorn & Hogarth, 1975). Unit weights are not estimated from the data and therefore do not have standard errors, as a result of which they can contribute to reduced estimation error as compared to a weighted average, especially when sample sizes and effect sizes are unequal (Bonett, 2008; Einhorn & Hogarth, 1975).

2.4. Analysis methods

2.4.1. Within-study analysis

In the within-study analysis, pairs of experimental conditions were categorised (e.g., no NDT vs. NDT, young participants vs. old participants, etc.). A meta-analysis was performed for a category when at least four studies were available in that category, following the recommendation by Fu et al. (2011). The 21 identified categories are shown in Fig. 3.

Because all studies used the same unit to measure TOTs (seconds), the meta-analyses were performed on the raw (unstandardized) difference between mean TOTs (D). The use of Ds allows for intuitive interpretations as compared to standardised effect size measures (Bond, Wiitala, & Richard, 2003; Higgins & Green, 2005). In other words, we described the effect of an independent variable in seconds instead of a dimensionless index such as Cohen's *d*. The use of seconds as a unit allows for easy interpretation in regard to practical applications (e.g., time budgets, look-ahead time of sensors) and the scientific literature in general (e.g., literature about brake response times, psychometric literature about reaction times).

The outcome of the meta-analysis was the unweighted average D per category. An absolute average D of 1 s was interpreted as a strong effect, and an absolute average D of 0.5 s was interpreted as a moderate effect.

In addition, we examined whether the D values differ from 0 (i.e., no effect), using a two-sided Wilcoxon signed-rank test with an alpha value of 0.05. This statistical test is conservative because a significant effect ($p < 0.05$) can only be obtained when six or more studies (D values) are available (i.e., $2 * 0.5^6 = 0.031 < 0.05$).

2.4.2. Between-study analysis

In the between-study analysis, we examined the correlations between 18 study variables (Table 1) and the mean TOTs. The 18 selected study variables were related to the driver, the automation system, the human-machine interface, the take-over situation, and the experimental set-up. More specifically, the variables concerned the mean age of participants, simulator fidelity, the level of automation, the modality of the TOR, the non-driving task (the modality of the task and if a device needed to be held in the hands), and the take-over situation (urgency of the scenario, complexity of the required driver response, and interaction with other road users). The selection of the study variables was based on the narrative reviews introduced above, studies providing guidelines for human factors research in the automated driving domain (e.g., De Winter, Happee, Martens, & Stanton, 2014; Gold, Naujoks, Radlmayr, Bellem, & Jarosch, 2018; Naujoks, Befelein, Wiedemann, & Neukum, 2017), and previous studies concerning driver response times in manual driving (e.g., Green, 2000; Summala, 2000). These variables were also selected based on whether they were available from the papers. For example, the physical intensity of the TOR, the level of drowsiness of the driver, and the duration of automated driving were not included as study variables, because these variables were often not documented, even though they are likely to affect the mean TOT.

We used Pearson product-moment correlations (equivalent to point-biserial correlations if the study variable is binary) and Spearman rank-order correlations to describe the relationships between the mean TOT and the study variables. The Spearman rank-order correlation is robust to tailed distributions and outliers.

A standard technique for assessing publication bias is to create a scatter plot showing the study outcome measures on the x-axis and a measure of sample size or precision on the y-axis, also called a funnel plot. An asymmetric relationship, where there exists a correlation between sample size and outcome measure, can be indicative of publication bias (Begg & Mazumdar, 1994; Egger, Smith, Schneider, & Minder, 1997; Deeks, Macaskill, & Irwig, 2005). We expected no effects of publication bias regarding the mean TOT, as high and low TOTs could be regarded as equally interesting to authors, publishers, and editors. Nonetheless, we assessed whether the mean TOT was correlated with the corresponding sample size.

Additionally, we computed correlations between sample size and all study variables (Table 1). These correlations allowed us to determine whether larger studies were associated with specific types of study design.

2.4.3. Linear mixed-effects model

We estimated a linear mixed-effects model describing the impact of the study variables on the mean TOTs, using the same dataset as the between-study analysis. A study-specific error term ϑ was introduced to capture unobserved effects that affect mean TOTs within a study (i.e., the intercept differs between studies).

The model was estimated using the 'Mixed Model' command in SPSS 24 (IBM Corporation, 2016) with the estimation method restricted maximum-likelihood (REML) (Molenberghs, Kenward, & Verbeke, 2009; Zuur, Leno, Walker, Saveliev, & Smith, 2009). Goodness-of-fit measures (log likelihood) and information criteria (AIC, BIC) were used to compare alternative model specifications. The SPSS script is provided in the supplementary materials.

A log-normal probability density function was found to fit the mean TOT distribution better than the normal probability density function. All variables listed in Table 1 were tested as potential explanatory factors. The variables included in the final specification were selected based on their meaning (i.e., we selected non-redundant variables) and statistical significance ($p < 0.05$). One parameter was associated with each level of the explanatory variables and differences between levels were tested by comparing alternative model specifications. Levels that did not differ significantly were merged. Variables that had a non-significant impact on the mean TOTs were excluded one by one. When a variable could not be extracted from one or more studies, a dummy variable was created to indicate the missing values. This variable was included in the model equation in addition to the original variable (dummy variable adjustment method). The advantage of this approach is that all observations could be analysed.

Table 1
Study variables and coding.

Study variable	Coding	Description
1. Age (AGE)	Years	Mean age of the participant group.
2. Level of automated driving (LAD)	0 = L2; 1 = L3 and above	The level of automated driving (SAE International, 2016) as reported by the authors of the paper. In L2 automated driving, participants are in charge of the monitoring task. In L3 automated driving and above, the drivers are not supposed to monitor the driving environment.
3. Simulator (SIM)[1]	0 = low fidelity	A desktop-based simulator or a simulator providing the environment through computer monitors.
	1 = medium fidelity	An instrumented-cabin simulator or a simulator providing more than 120 deg horizontal field of view.
	2 = high fidelity	A simulator with motion platform or a real car.
4. Visual TOR (TOR_V)	0 = no; 1 = yes	Whether the TOR contains a visual stimulus.
5. Auditory TOR (TOR_A)	0 = no; 1 = yes	Whether the TOR contains an auditory stimulus. We applied no differentiation between vocal and acoustic TORs.
6. Vibrotactile TOR (TOR_VT)	0 = no; 1 = yes	Whether the TOR contains a vibrotactile stimulus (i.e., vibrations are provided on one or more locations on the human body).
7. Presence of TOR (TOR_P)	0 = no; 1 = yes	Whether a TOR is implemented.
8. Visual NDT (NDT_V)	0 = no; 1 yes	Whether the NDT is visual (e.g., reading, watching a movie).
9. Auditory NDT (NDT_A)	0 = no; 1 = yes	Whether the NDT is auditory (e.g., listening to the radio, watching movies with sound, communicating with the instructors and answering questions verbally).
10. Motoric NDT (NDT_M)	0 = no; 1 = yes	Whether dynamic operation by hand is needed to perform the NDT (e.g., texting and tapping).
11. Cognitive load (NDT_C)	0 = normal cognitive load; 1 = high cognitive load	NDTs that require working memory (e.g., N-back task) were assigned to the high cognitive load category. Otherwise, the task was assumed to involve normal cognitive load.
12. Hand holding a device (HAND)	0 = hands-free (No non-driving task, the non-driving task does not require a device, or the non-driving task is performed using a fixed device); 1 = handheld	Whether a device is handheld when undertaking the non-driving task.
13. Presence of a non-driving task (NDT_P)	0 = no; 1 = yes	Whether a non-driving task is performed.
14. Time budget to collision (TBTC)[1]	Ratio variable	The available time budget for a response from the initiation of the take-over stimulus until the collision with an obstacle.
15. Time budget to other boundaries (TBTB)[1]	Ratio variable	The time from the initiation of the take-over stimulus until reaching the boundaries of the automated driving system other than collisions (e.g., due to the end of the automated zone, missing lane markers, or system failure).
16. Urgency (URG)[1]	0 = low urgency	No foreseeable collision risk or a high time budget to collision ($\geq$15 s).
	1 = medium urgency	Potential collision risk or disturbance to other road users if no response was made (e.g., the ego car drifting to the adjacent lane containing traffic), or a medium time budget (between 8 s and 15 s).
	2 = high urgency	Immediate risk of collision (time budget $\leq$ 8 s) if no response was made, or the participants were instructed to react to a stimulus as quickly as possible.
17. Driver response (DRE)[1]	1 = low complexity	The participant had to take over control on a straight road by stabilising the vehicle in its lane.
	2 = medium complexity	The take-over scenarios required a specific driver response (braking or steering), such as when encountering a road narrowing, road constructions, or decelerating vehicles ahead, or when having to take over control on a curvy road.

Table 1 (*continued*)

Study variable	Coding	Description
	3 = high complexity	The take-over scenario requires complex driver decision making. The participant had to decide whether to brake or steer in response to the event.
18. Interaction with other road users (IRU)	0 = no	There were no other road users around, or other road users could not affect the driver's decision-making.
	1 = yes	The participants had to take into account one or more other road users when choosing their optimal take-over action. For example, participants had to take over control while driving in the middle lane while the left lane contained traffic.

[1] *Note*. The fidelity of the simulator was identified according to De Winter, Happee et al. (2014). The classification of URG and DRE was adapted from Gold, Naujoks et al. (2018). URG combines the TBTC and TBTB variables. TOR = Take-over request, NDT = Non-driving task.

3. Results

3.1. Study characteristics

The 129 included studies yielded 520 mean TOT observations from 4556 participants. 45 studies were conducted in a high fidelity driving simulator with motion platform (40 studies) or in a real car (5 studies). The 129 studies comprised 68 papers from conference proceedings, 40 journals articles, 3 technical reports, 16 chapters from a PhD or master thesis, and 2 posters. The year of publication ranged between 2000 and 2018, with the majority of the studies being published in and after 2015 (116 out of 129). A list of the included studies, the scores per study variable, and a MATLAB script that processes these data are provided in the supplementary materials.

The mean TOT across studies and conditions ranged from 0.69 s to 19.79 s, and the average mean TOT was 2.72 s (SD = 1.45, n = 520). Fig. 2 shows the distribution of the mean TOTs, which is right-skewed.

3.2. Within-study analysis

Twenty-one categories with four or more studies were identified in the experimental analysis. Fig. 3 shows the average difference in mean TOTs (D) for each category (squares), as well as the Ds from each study (smaller circles). The presence of non-driving tasks and the modality of the take-over request (TOR) were frequently used independent variables.

The following statistically significant findings stand out from Fig. 3:

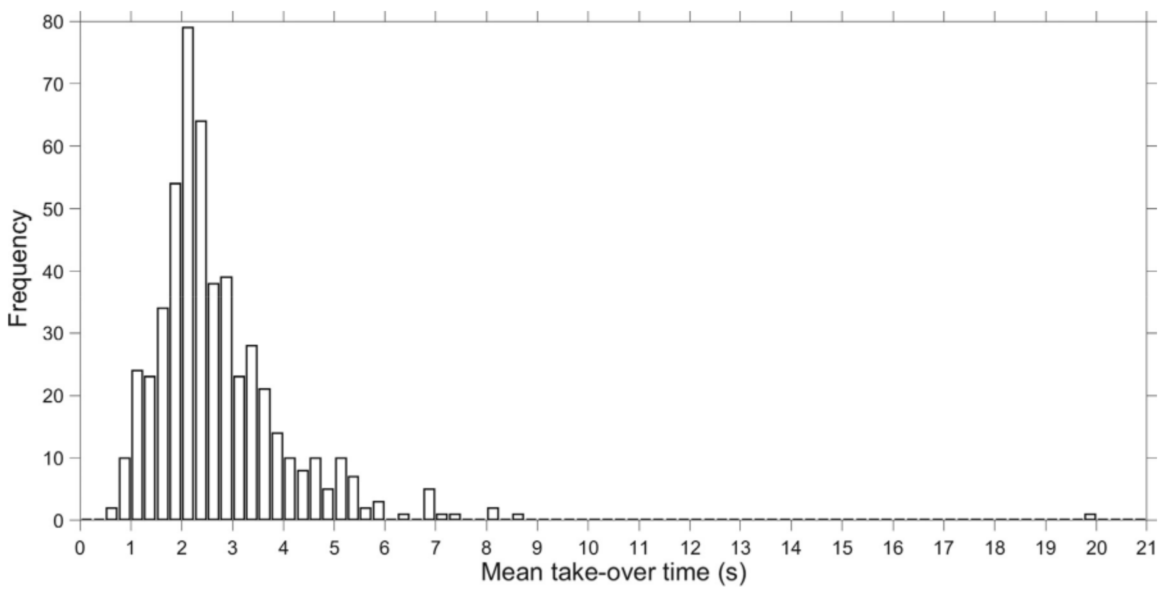

Fig. 2. Distribution of 520 mean take-over times (TOTs) reported in the included studies. The bin width is 0.25 s.

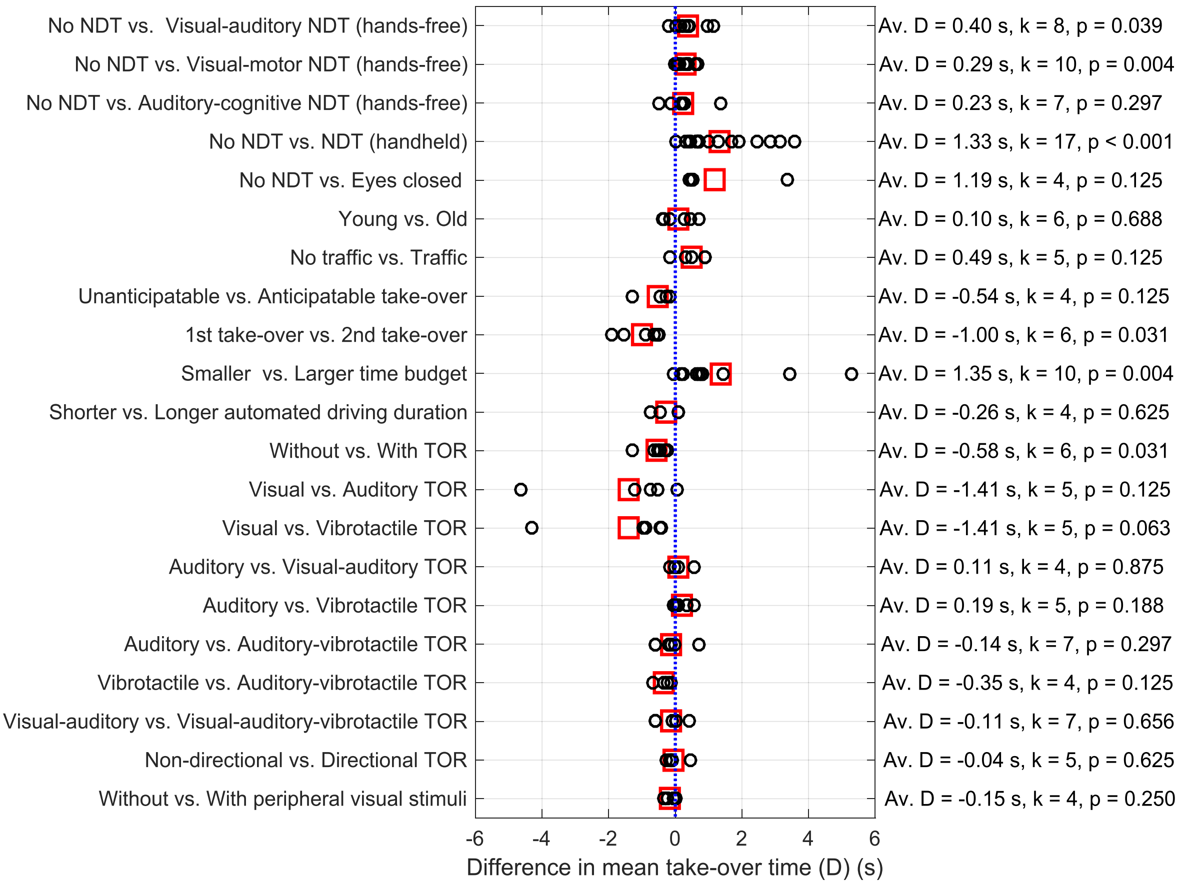

Fig. 3. Within-study effects. The circles represent the difference in mean take-over time between two conditions (D) of a particular study. A positive D indicates that a larger mean TOT was obtained from the latter condition compared to the former condition. The large square markers represent the average D in the category. k represents the number of studies (D values) in the category. p is the p-value from a two-sided signed rank test for the hypothesis that the D values come from a distribution having a median of 0. TOR = Take-over request, NDT = Non-driving task. A directional TOR is a TOR that is informative about the location of the hazard. Peripheral visual stimuli are stimuli that indicate the status of the automation or the environment (e.g., using ambient LEDs).

- A strong effect of time budget was found, with a higher mean TOT (average D = 1.35 s) for a large time budget compared to a small time budget.
- The mean TOT was substantially lower when taking over control for the second time (when asked to take over twice in the same driving session or perform the same scenario in a second driving session) compared to the first time (average D = −1.00 s).
- The use of a handheld device strongly increased the mean TOT (average D = 1.33 s).
- For hands-free non-driving tasks, performing a visual non-driving task slightly increased the mean TOT compared to not performing a non-driving task (average D = 0.29 s).
- The presentation of a TOR moderately decreased response times compared to when no TOR was provided (average D = −0.58 s).

In addition, the following findings are noteworthy, although based on five or fewer studies.

- Having eyes closed before taking over control strongly increased TOTs compared to not performing a non-driving task and staying alert (average D = 1.19 s).
- Strongly reduced TOTs were found when an auditory or vibrotactile TOR was provided compared to a visual-only TOR (average D = −1.41 and −1.41 s, respectively).
- Being able to anticipate the TOR (e.g., when the TOR was periodically scheduled or could be anticipated from environmental cues such as the traffic and weather) moderately shortened the mean TOT (average D = −0.54 s).
- The effect of traffic compared to no traffic was moderate (average D = 0.49 s).

3.3. Between-study analysis

The correlations between the 18 study variables and the mean TOT are shown in Table 2. As in the within-study analysis, urgency of the take-over scenario and holding a handheld device showed strong correlations with the mean TOT: higher-urgency levels and shorter time budgets were associated with lower mean TOT (URG: $r = -0.44$; $\rho = -0.42$; TBTC: $r = 0.53$, $\rho = 0.43$; TBTB: $r = 0.73$, $\rho = 0.31$), and holding a handheld device (HAND) yielded higher mean TOT ($r = 0.30$, $\rho = 0.35$). The correlations between mean TOT and other categorical variables were weak to moderate, with absolute values below 0.3. That is, the mean TOT did not substantially correlate with the modality of the TOR or the type of non-driving task.

Additionally, we calculated the correlation between the mean and the standard deviation of the TOTs and found a strong association (Fig. 4; $r = 0.82$; $\rho = 0.73$, $n = 397$). The correlations between the mean TOTs and the three continuous study variables: AGE, TBTC, and TBTB are depicted in Fig. 5.

A weak to moderate correlation was observed between sample size and mean TOT ($r = 0.21$, $\rho = 0.14$, $n = 520$), see Fig. 6. The correlations between sample size and all study variables were weak with an absolute r and ρ smaller than 0.20, except for the correlation with time budget to collision (TBTC, $r = 0.26$, $\rho = 0.27$, $n = 240$) and simulator fidelity where ρ was smaller than -0.20 (SIM; $r = -0.13$, $\rho = -0.25$, $n = 520$). In other words, studies in high-fidelity simulators involved smaller sample sizes than studies in low-fidelity simulators.

Table 3 shows the correlations between the study variables, providing insight into the patterns of the experimental design. For a higher level of automation, the studies tended to implement a TOR ($\rho = 0.65$), instruct the participants to perform a non-driving task ($\rho = 0.34$), and provide longer time budgets to collision ($\rho = 0.35$). These observations are in accordance with the definition of SAE International (2016).

Concerning the modalities of the non-driving task and TOR, strong positive correlations were found between the presence of a motoric and visual non-driving task ($\rho = 0.68$). Visual and auditory TORs tended to be combined ($\rho = 0.35$), which was not the case for the auditory and vibrotactile modalities ($\rho = -0.35$).

Table 2
Correlations between the study variables and mean TOTs in the between-study analysis. Pearson product-moment correlations (r) and Spearman rank-order correlations (ρ) are both reported.

Study variable	Correlation with mean TOT					
	r	ρ	Study variable conditions	*n*	Average mean TOT (s)	*SD* mean TOT (s)
1. AGE	0.22	0.24	–	485	–	–
2. SIM	−0.04	0.02	0 (low fidelity)	81	2.67	2.39
			1 (medium fidelity)	268	2.84	1.33
			2 (high fidelity)	171	2.57	0.95
3. LAD	0.15	0.19	0 (L2)	62	2.14	1.17
			1 (L3 and above)	458	2.80	1.47
4. TOR_V	0.04	0.08	0 (no visual TOR)	160	2.63	1.75
			1 (with visual TOR)	360	2.77	1.29
5. TOR_A	0.12	0.12	0 (no auditory TOR)	84	2.33	1.00
			1 (with auditory TOR)	436	2.80	1.51
6. TOR_VT	−0.11	−0.10	0 (no vibrotactile TOR)	447	2.79	1.50
			1 (with vibrotactile TOR)	73	2.35	0.99
7. TOR_P	0.06	0.06	0 (no TOR)	34	2.40	1.08
			1 (TOR present)	486	2.75	1.47
8. NDT_V	0.13	0.13	0 (no visual NDT)	204	2.49	1.12
			1 (the NDT is visual)	309	2.89	1.62
9. NDT_A	−0.03	−0.07	0 (no auditory NDT)	384	2.75	1.29
			1 (the NDT is auditory)	129	2.67	1.88
10. NDT_M	−0.01	−0.04	0 (no motoric NDT)	289	2.74	1.27
			1 (the NDT requires a motoric manoeuvre)	224	2.72	1.67
11. NDT_C	−0.05	−0.11	0 (without highly cognitively demanding NDT)	385	2.78	1.28
			1 (with highly cognitively demanding NDT)	128	2.60	1.90
12. HAND	0.30	0.35	0 (no handheld device)	371	2.54	1.42
			1 (NDT device held in the hands)	108	3.61	1.46
13. NDT_P	0.11	0.11	0 (no NDT present at the moment of TOR)	143	2.46	1.17
			1 (NDT present at the moment of TOR)	377	2.82	1.53
14. URG	−0.44	−0.42	0 (low urgency)	83	3.95	2.35
			1 (medium urgency)	114	3.03	1.32
			2 (high urgency)	295	2.25	0.81
15. DRE	−0.16	−0.07	0 (low response complexity)	108	3.43	2.21
			1 (medium response complexity)	134	2.34	1.16
			2 (high response complexity)	253	2.66	1.04
16. IRU	0.08	0.14	0 (no interaction with other road users)	344	2.67	1.55
			1 (interaction with other road users)	141	2.93	1.24
17. TBTC	0.53	0.43	—	240	—	—
18. TBTB	0.73	0.31	—	160	—	—

Note. TOR = Take-over request, NDT = Non-driving task.

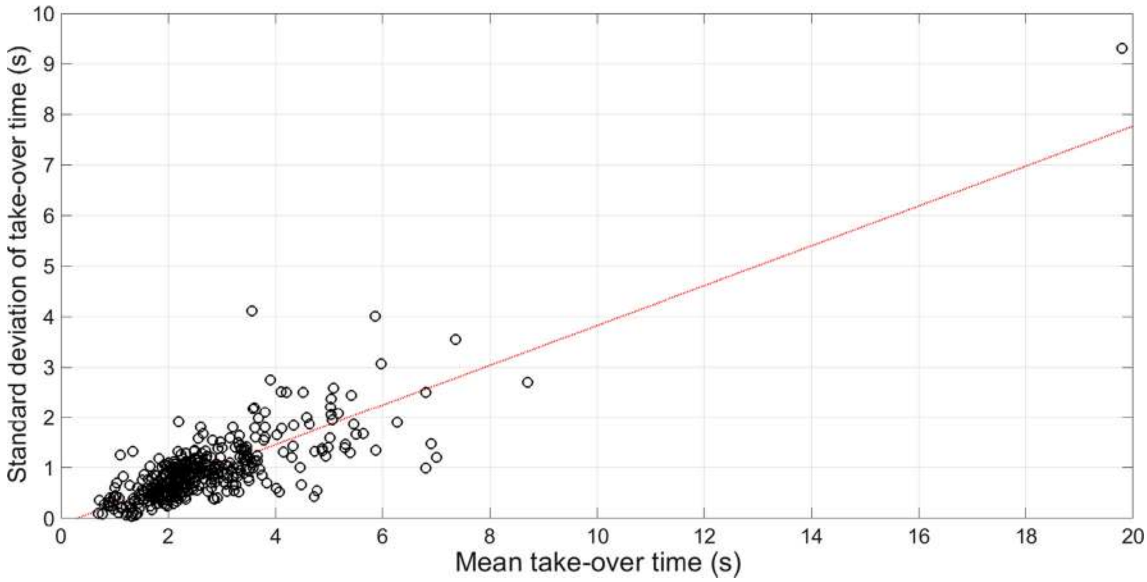

Fig. 4. Scatter plot of the standard deviation of the take-over time (*SD* TOT) as a function of the mean take-over time (mean TOT), with a fitted least-squares regression line (n = 397).

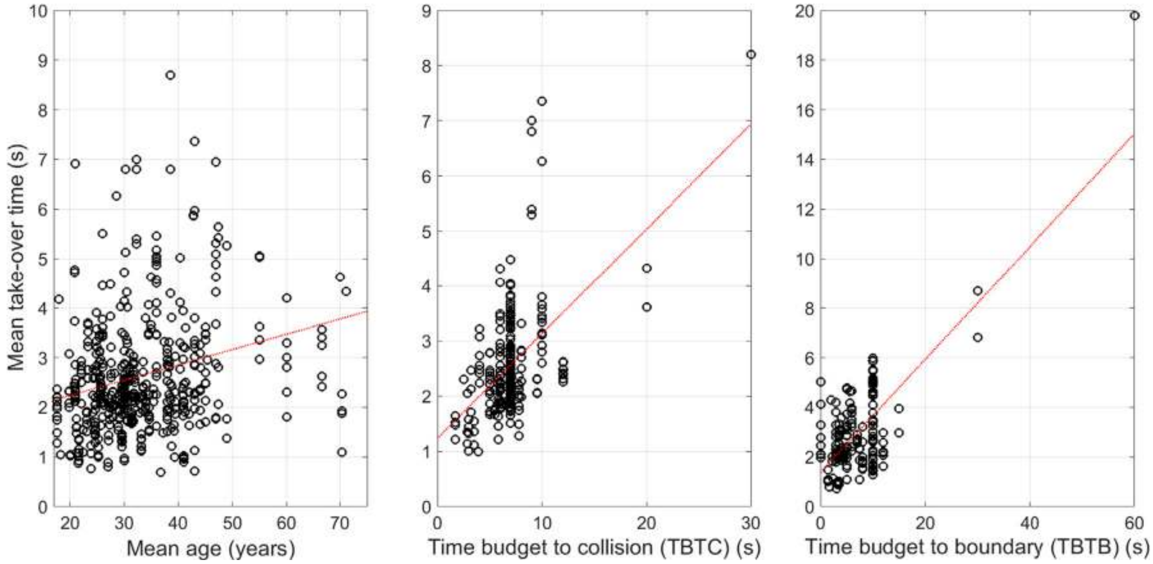

Fig. 5. Scatter plot of the mean take-over time (mean TOT) as a function of (a) mean age of the participant group (n = 485), (b) take-over time budget to collision (TBTC, n = 240), (c) take-over time budget to other boundaries (TBTB, n = 160), with a fitted least-squares regression line. A TBTB of 0 s means that the automation was deactivated at the moment of the take-over stimulus.

A complex driver response was more likely to be required when the take-over situation was more urgent (ρ = 0.52), and when other road users were involved in the take-over process (ρ = 0.46). Also, it is worth noting that studies that used higher fidelity simulators tended to employ older participants (ρ = 0.38).

3.4. Linear mixed-effects model

The goodness of fit indicators of the linear mixed-effects model are shown in Table 4. Table 5 shows the estimation results where effects for most study variables were strongly statistically significant (i.e., low p-values). The model predicting ln(TOT) (i.e., the natural logarithm of the mean TOT) is specified according to Eq. (1):

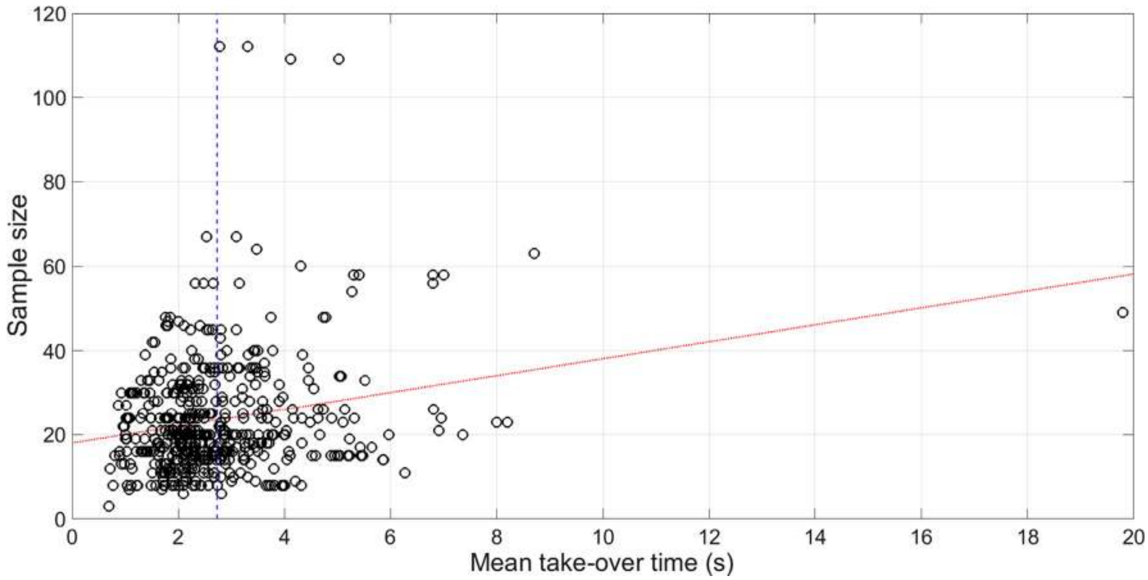

Fig. 6. Scatter plot of the sample size of the participant group as a function of the corresponding mean take-over time (mean TOT), with a fitted least-squares regression line (n = 520). The vertical line represents the grand mean TOT.

$$\begin{aligned} \ln(\mathrm{TOT}) = \alpha + \beta_{\mathrm{LAD}} \cdot \mathrm{LAD} + \beta_{\mathrm{TOR_A}} \cdot \mathrm{TOR_A} + \beta_{\mathrm{TOR_{VT}}} \cdot \mathrm{TOR_{VT}} \\ + \beta_{\mathrm{NDT_V}} \cdot \mathrm{NDT_V} + \beta_{\mathrm{MissNDT_V}} \cdot \mathrm{MissNDT_V} + \beta_{\mathrm{Hand}} \cdot \mathrm{Hand} + \beta_{\mathrm{MissHand}} \cdot \mathrm{MissHand} \\ + \beta_{\mathrm{URG_{High}}} \cdot \mathrm{URG_{High}} + \beta_{\mathrm{URG_{Med}}} \cdot \mathrm{URG_{Med}} + \beta_{\mathrm{MissURG}} \cdot \mathrm{MissURG} \\ + \beta_{\mathrm{IRU}} \cdot \mathrm{IRU} + \beta_{\mathrm{MissIRU}} \cdot \mathrm{MissIRU} + \gamma \cdot \vartheta + \sigma \cdot \varepsilon \end{aligned} \quad (1)$$

where α is the intercept, $\beta_{\mathrm{LAD}}, \beta_{\mathrm{TOR_A}}, \beta_{\mathrm{TOR_{VT}}}, \beta_{\mathrm{NDT_V}}, \beta_{\mathrm{Hand}}, \beta_{\mathrm{URG_{High}}}, \beta_{\mathrm{URG_{Med}}}, \beta_{\mathrm{IRU}}$ are the parameters associated with the study variables listed in Table 5, $\beta_{\mathrm{MissNDT_V}}, \beta_{\mathrm{MissHand}}, \beta_{\mathrm{MissURG}}, \beta_{\mathrm{MissIRU}}$ are parameters associated with the dummy variables indicating the missing values, γ is the parameter associated with the study-specific error term $\vartheta \sim \mathrm{N}(0,1)$, and σ is the parameter associated with the observation-specific error term $\varepsilon \sim \mathrm{N}(0,1)$. The study-specific error term captures between-study variance and the observation-specific error term captures residual variance between observations. We selected the study variables based on statistical significance. The other study variables listed in Table 1 did not significantly influence the mean TOT.

Studies carried out with high levels of automation (≥SAE level 3) showed longer mean TOTs than studies with partial automation (SAE Level 2). The fidelity of the driving simulator did not significantly influence mean TOTs (i.e., was not a sufficiently strong predictor to be included in the model as a predictor variable). Auditory and vibrotactile TOTs were associated with shorter mean TOTs, whereas visual warnings did not have a statistically significant impact on mean TOTs. Participants were slower in taking over control when they were engaged in visual non-driving tasks. Other types of non-driving tasks (auditory, motoric, and cognitive) did not significantly influence mean TOTs. Participants showed longer mean TOTs when holding a handheld device. Take-over situations with high and medium levels of urgency were related to shorter TOTs than situations with a low level of urgency. Finally, interacting with one or more other road users was associated with a longer mean TOT as compared to when no road users were driving in the vicinity.

The model coefficients in Table 5 are defined on a logarithmic scale, which enhances the model fit but complicates the interpretation. To illustrate the impact of the study variables on the mean TOT in seconds, we used the linear mixed-effects model to calculate the mean TOT in a baseline observation and the mean TOT where one variable was changed while keeping all the other variables fixed. In the baseline observation, the level of automation was high (≥SAE level 3), the TOR was auditory, the NDT was visual, and the level of urgency was high. In addition, drivers did not use a handheld device and did not have to take into account other road users. These baseline values were selected because they represent the majority of the conditions available. The impact of the study variables on the mean TOTs is shown in Table 6. The level of urgency, holding a device in the hands, and the use of an auditory TOR had the largest impact on the mean TOTs.

4. Discussion

4.1. Findings from this meta-analysis

This meta-analysis quantified the determinants of mean take-over time (TOT) as observed in 129 experiments using three complementary approaches: a within-study analysis, a between-study analysis, and a linear mixed-effects model. The

Table 3
Spearman rank-order correlations (ρ) between the predictor variables shown in Table 2.

	1. AGE	2. SIM	3. LAD	4. TOR_V	5. TOR_A	6. TOR_VT	7. TOR_P	8. NDT_V	9. NDT_A	10. NDT_M	11. NDT_C	12. HAND	13. NDT_P	14. URG	15. DRE	16. IRU	17. TBTC	18. TBTB
1.AGE	–																	
2.SIM	0.38	–																
3.LAD	−0.07	0.03	–															
4.TOR_V	0.07	0.26	0.24	–														
5.TOR_A	0.18	0.18	0.34	0.35	–													
6.TOR_VT	−0.21	−0.30	0.13	−0.13	−0.35	–												
7.TOR_P	0.05	0.08	0.65	0.40	0.60	0.11	–											
8.NDT_V	−0.12	−0.08	0.30	0.12	0.08	0.10	0.19	–										
9.NDT_A	0.04	−0.17	0.08	−0.09	−0.04	0.07	0.04	0.03	–									
10.NDT_M	−0.18	−0.01	0.28	0.13	0.07	0.13	0.18	0.68	−0.21	–								
11.NDT_C	−0.09	−0.21	0.16	−0.06	0.01	0.09	0.10	0.13	0.27	0.29	–							
12.HAND	−0.05	0.03	0.21	0.09	0.10	0.04	0.11	0.31	−0.14	0.24	−0.12	–						
13.NDT_P	−0.08	−0.08	0.34	0.08	0.08	0.11	0.22	0.77	0.36	0.55	0.33	0.35	–					
14.URG	−0.27	0.01	−0.16	−0.04	−0.14	0.10	−0.13	−0.07	−0.02	0.12	0.13	−0.18	−0.06	–				
15.DRE	−0.10	0.05	0.06	0.00	−0.01	0.09	−0.01	0.05	−0.04	0.17	0.10	0.00	0.09	0.52	–			
16.IRU	0.12	0.03	0.08	0.03	0.01	−0.06	0.00	0.01	−0.07	0.05	0.16	0.04	0.02	0.10	0.46	–		
17.TBTC	−0.17	−0.30	0.35	0.13	0.02	0.20	0.22	0.09	0.01	0.02	−0.03	0.21	0.17	−0.61	−0.12	−0.01	–	
18.TBTB	−0.24	0.08	0.22	−0.19	−0.12	0.13	0.01	0.12	−0.03	0.21	−0.06	0.28	0.16	−0.33	0.03	−0.01	N/A	–

Table 4
Characteristics of the linear mixed-effects model.

Number of studies	129
Number of observations	520
−2 Restricted log likelihood	132.2
Akaike's Information Criterion (AIC)	136.2
Schwarz's Bayesian Criterion (BIC)	144.6

Table 5
Results of the linear mixed-effects model. The parameters represent the unconditional marginal effects of the study variables on the natural logarithm of the mean TOT.

Variable	Description	Parameter	Estimate	df	t	p
	Intercept	α	1.0620	203.39	11.12	**$9.39 * 10^{-23}$**
LAD	Equal to 1 when the level of automated driving (SAE International, 2016) is L3 or above as reported by the authors of the paper.	β_{LAD}	0.2850	140.51	3.11	**$2.24 * 10^{-3}$**
TOR_A	Equal to 1 when the TOR contains an auditory stimulus.	β_{TOR_A}	−0.2129	451.50	−5.29	**$1.89 * 10^{-7}$**
TOR_{VT}	Equal to 1 when the TOR contains a vibrotactile stimulus.	$\beta_{TOR_{VT}}$	−0.1801	450.54	−4.00	**$7.53 * 10^{-5}$**
NDT_V	Equal to 1 when the NDT is visual.	β_{NDT_V}	0.0975	450.68	3.32	**$9.56 * 10^{-4}$**
$MissNDT_V$	Equal to 1 when it is not mentioned whether the NDT is visual.	$\beta_{MissNDT_V}$	−0.1545	151.99	−0.62	$5.36 * 10^{-1}$
HAND	Equal to 1 when a device is handheld when undertaking the NDT.	β_{HAND}	0.2310	461.09	6.32	**$6.22 * 10^{-10}$**
MissHAND	Equal to 1 when it is not mentioned whether a device is handheld when undertaking the NDT.	$\beta_{MissHAND}$	−0.0559	507.00	−0.63	$5.29 * 10^{-1}$
URG_{High}	Equal to 1 when there is an immediate risk of collision (time budget ≤ 8 s), or the participants were instructed to react to a stimulus as quickly as possible.	$\beta_{URG_{High}}$	−0.4658	491.88	−8.14	**$3.17 * 10^{-15}$**
URG_{Med}	Equal to 1 when there is potential collision risk or disturbance to other road users if no response is made, or a medium time budget (between 8 s and 15 s).	$\beta_{URG_{Med}}$	−0.2193	506.95	−3.62	**$3.21 * 10^{-4}$**
MissURG	Equal to 1 when it is not mentioned whether there is an immediate risk.	$\beta_{MissURG}$	−0.1417	145.68	−0.77	$4.44 * 10^{-1}$
IRU	Equal to 1 when the participants had to take into account one or more other road users when choosing their optimal take-over action.	β_{IRU}	0.1821	502.89	4.57	**$6.13 * 10^{-6}$**
MissIRU	Equal to 1 when it is not mentioned whether the driver had to take into account one or more other road users when choosing their optimal take-over action.	$\beta_{MissIRU}$	−0.0220	336.49	−0.18	$8.59 * 10^{-1}$
Error term	**Description**	**Parameter**	**Estimate**		**Wald-Z**	**p**
ϑ_s	Study-specific error term (between-study variance)	γ	0.1357		6.89	**$5.40 * 10^{-12}$**
ε_S	Observation-specific error term (between-observation variance)	σ	0.0375		13.73	**$6.73 * 10^{-43}$**

$p < 0.05$ is indicated in boldface.

Table 6
Effect of the study variables on the baseline TOTs (average baseline TOT = 2.15 s).

Variable	Estimated mean TOT (s)
LAD = 0	1.62
TOR_A = 0	2.66
TOR_{VT} = 1	1.80
NDT_V = 0	1.95
Hand = 1	2.71
URG_{Low} = 1	3.43
URG_{Med} = 1	2.75
IRU = 1	2.58

within-study analysis provides a synthesis of causal experimental effects (Fig. 3). The between-study analysis is based on correlations that involve hundreds of mean TOT values (Table 2), and may, therefore, feature higher generalizability than the within-study analysis. However, the between-study analysis is a synthesis of correlations rather than causal effects, and may therefore be susceptible to various confounding factors. The linear mixed-effects model is statistically powerful because it uses the dataset of the between-study analysis while taking into account whether the mean TOT values were obtained from the same study. Although all models are wrong if taken literally (Box, 1976), we would argue that our three-fold complementary approach provides a good picture of the current take-over literature.

Several main findings stand out from the three meta-analysis methods. First, the urgency of the situation, defined in terms of (1) the hand-coded urgency level (URG), (2) time budget to collision with an obstacle (TBTC), and (3) time budget to boundaries (TBTB), has substantial associations with the mean TOT. In other words, if more time is available, drivers use more time to take over. This observation is consistent with previous reviews (see Section 1.2) and can be interpreted using a literature review by Summala (2000) on brake reaction times in manual driving. Summala argued that a distinction exists between drivers' ability to intervene quickly and their motivation to intervene, and explained that "it is not always necessary to react as soon as possible". If there is sufficient time, drivers do not take-over as quickly as they can, but first assess the situation (e.g., by checking the mirrors) (Gold, Damböck et al., 2013) and resume an optimal driving posture (e.g., by adjusting the seating position) (Zhang et al., in press) before taking over.

The second finding is that performing a non-driving task with a handheld device strongly increases the mean TOT, as confirmed by each of the three analyses. Among the studies without a handheld device, performing a visual NDT yielded a moderate increase in mean TOT as compared to not performing such a task. The mixed-effects model confirmed that engagement in a visual NDT increased mean TOTs. The other modalities of the NDT, that is, whether the task demand is auditory or cognitive, did not show significant associations with the mean TOTs. Zhang et al. (in press) examined driver perception and movement response times during take-over process, and found that physically switching arm posture from the current NDT to the driving control task requires more time than perceiving and cognitively processing the take-over stimuli, especially when the arm movement amplitude is high (cf. Fitts, 1954).

Third, a high level of automation (SAE L3 and above) showed higher mean TOTs compared to partial automation (SAE L2), possibly due to a combined effect of a longer time budget, lower urgency, and more involvement in (handheld) NDTs, consistent with the definition provided by the SAE International (2016).

Fourth, as shown in the within-study analysis, prior experience with taking over has a strong effect: drivers responded about 1 s faster if the take-over scenario occurred the second time compared to the first time. The majority of the included studies (92 out of 129) used a within-subject design. In a review about brake response times in manual driving, Green (2000) pointed out that in most studies, participants performed multiple trials to generate more data. The repeated trials would contribute to shorter response times, which calls for caution when interpreting the results. Our meta-analysis also showed that drivers responded about 0.5 s faster when the TOR could be anticipated from task-related or environmental cues. This finding is in line with publications indicating that expectancy is an important factor influencing brake response times (Green, 2000; Warshawsky-Livne & Shinar, 2002; Young & Stanton, 2007).

Fifth, visual-only TORs showed longer mean TOTs than auditory or vibrotactile TORs. The mixed-effects model further showed that auditory and vibrotactile TORs reduce the mean TOTs as compared to when such TORs are not present or the TOR is visual only. Petermeijer, Doubek, and De Winter (2017) argued that a visual-only warning is not suitable as a TOR, as drivers may overlook a visual signal (especially if they are performing a visually distracting NDT) or may not interpret a visual signal as urgent. Auditory warnings, on the other hand, are well established due to their omnidirectional characteristics (Bazilinskyy & De Winter, 2015). Vibrotactile TORs are effective as well, as they can attract the driver's attention when the driver is performing a visual or auditory NDT (Petermeijer et al., 2016).

Sixth, we found no clear effect of age in the within-study analysis or the multi-level model, which is interesting because age is known to be associated with a slower speed of processing (e.g., Salthouse, 2009). One possible explanation is that TOTs largely reflect motivational processes, not biological limitations, as pointed out above. For example, although older drivers have a slower simple reaction time, they could have a more cautious driving style and are likely to take over quickly even when not strictly necessary (Körber, Gold, Lechner, & Bengler, 2016). Compensatory behaviours, such as a less intensive involvement in NDTs, may also alleviate ageing effects (Clark & Feng, 2017). Furthermore, the positive correlation between age and mean TOT in the between-study analysis may point to a confounding effect, where older drivers have participated in different types of experiments, as discussed below. We argue that the lack of observed age effects in the within-study analysis is not due to range restrictions, as the differences in mean age for the six included studies were substantial (23 vs. 67 years, 20 vs. 70 years, 34 vs. 60 years, 18 vs 70 years, 18 vs. 37 years, 26 vs. 71 years). It has been recommended that future take-over studies include even older drivers, above 80 years of age (e.g., Körber et al., 2016; Li, Blythe, Guo, & Namdeo, 2018). Additionally, we would recommend that future research on the effect of biological age on TOT should try to obtain a more in-depth understanding by examining the effects of covariates, such as years of driving experience, psychometric performance (e.g., simple, reaction time, perceptual speed), and sensation seeking scores.

Finally, we found a moderate effect of surrounding traffic (IRU) on the mean TOT. This can be explained by the fact that drivers, in case of surrounding traffic, need time for visual scanning and situation assessment before taking over. However, we found that a more complex driver response (i.e., higher DRE) was associated with a shorter mean TOT in the between-study analysis. This counterintuitive finding could be due to the strong positive correlation between DRE and urgency. In other words, although complex responses require cognitive processing time (see e.g., Gold, Naujoks, et al., 2018, and Green, 2000 for discussion in manual driving context), such responses are more likely to be performed in urgent situations, which are associated with lower mean TOTs.

4.2. Limitations

The current study was performed using mean TOTs. In the end, collision risk is not determined by the mean TOT, but by outliers in the TOT distribution (Horrey & Wickens, 2007). We found that the mean and standard deviation of TOT are highly correlated (r = 0.82; ρ = 0.73), indicating that mean TOTs are informative about the tail of the TOT distribution. However, we note that accidents may be due to extreme values (e.g., TOTs that exceed the 99.999th percentile, such as due to a driver being asleep behind the wheel). Our meta-analysis is not suitable for making inferences about crash likelihood or for proposing generic guidelines about what time budget constitutes safety.

A second limitation is that this meta-analysis investigated take-over *time*, not take-over *quality*. A number of studies found fast but also hazardous responses (severe braking or steering) under higher mental workload and short time budgets (e.g., Gold, Damböck et al., 2013; Clark & Feng, 2017; Ito, Takata, & Oosawa, 2016). Put differently, a short TOT does not necessarily indicate a safe situation, but could actually be a sign of hazard, because short TOTs typically occur in urgent situations for which an evasive manoeuvre may be needed. We found that directional TORs (i.e., TORs that are informative about the location of the hazard) have no beneficial effects on mean TOT, but this does not imply that directional TORs are ineffective, as they could be useful to enhance take-over quality (e.g., to enhance situation awareness). We recommend that researchers publish not only the mean and standard deviation of the TOT, but also provide data files with TOT values per event. This would allow meta-analysts to make inferences about the TOT distribution. Additional response variables, such as minimum time to collision and maximum longitudinal/lateral acceleration would enable the assessment of take-over quality.

Third, the starting moment of the take-over response is a source of ambiguity (Liu & Green, 2017). While some researchers used criteria such as "the moment the driver gave an input either on the pedals or the steering wheel" (Payre, Cestac, Dang, Vienne, & Delhomme, 2017), other researchers provided exact criteria. For example, Gold, Damböck et al. (2013) adopted a 2 degree steering angle or 10% brake pedal position, which was employed in a number of subsequent studies (Feldhütter, Gold, Schneider, & Bengler, 2017; Gold et al., 2015; Gold, Körber, Lechner, & Bengler, 2016; Gold et al., 2014; Gold, Lorenz, Damböck, & Bengler, 2013; Kerschbaum, Lorenz, & Bengler, 2015; Körber et al., 2016; Radlmayr et al., 2014). Somewhat different criteria can be found in other studies, such as absolute steering acceleration larger than 5 deg/s^2 (Zeeb, Buchner, & Schrauf, 2015, 2016). Another issue is that the TBTC posed an upper limit to the TOTs that could be observed; if a participant would not react at all (which sometimes happened, see Gold, Lorenz et al., 2013; Young & Stanton, 2007), their results were not taken into account in the reported mean TOT, which would underestimate the mean TOT.

Fourth, although a large number of study variables were investigated, there may still be unobserved study variables that affect TOT. Examples of hidden moderators are driving speed, the intensity of the TOR, and the state of the operator (e.g., whether he or she is fatigued or impaired by alcohol, see Wiedemann et al., 2018). Hidden moderators may also explain why the study-specific error term is strong in the mixed-effects model.

Fifth, nearly all included studies were conducted in a driving simulator. Despite advantages such as controllability and safety, driving simulators have limited fidelity, which raises the issue about behavioural validity (De Winter, Van Leeuwen, & Happee, 2012; Green, 2000; Kaptein, Theeuwes, & Van der Horst, 1996; Riener, 2010; Risto & Martens, 2014). Also, the driver's level of perceived risk perception may be low in simulators as compared to on-road conditions (Carsten & Jamson, 2011), which could discourage a fast take-over response. While the TOTs measured in simulator studies may not accurately reflect the numeric values of the TOTs in the real world, the results may still be valid concerning the direction of the effects (Kaptein et al., 1996).

Sixth, although our meta-analysis is much more comprehensive compared to previous reviews on the same topic (see Introduction), the within-study analysis would still benefit from a larger sample of studies. The Ds in Fig. 3 are based on 4 to 17 studies. Although each of these individual studies may present credible findings, more studies should be conducted to examine whether the experimental effects are generalizable.

Finally, as in any meta-analysis, there may be sources of bias or confounding effects. In a previous review on automated driving, De Winter, Happee et al. (2014) observed a confounder, namely that young participants are overrepresented in lower-fidelity driving simulators. The authors explained that lower-fidelity simulators are available at universities where participants are usually students, whereas companies with high-fidelity simulators tend to recruit middle-age drivers. A similar association between simulator fidelity level and participant age was observed in the between-study meta-analysis (ρ = 0.38). We also found studies in lower-fidelity simulators involved larger sample sizes (ρ = −0.25), which could be explained because students participate in relatively large amounts, e.g., for course credit. Such confounds affect some of the correlations in the between-study analysis (Table 2) but are controlled for in the mixed-effects model. In the between-study analysis, a weak-to-moderate correlation was observed between sample size and mean TOT, which may be related to the confounding effect of simulator fidelity discussed above.

We expect that publication bias regarding mean TOT in the between-study analysis is small as compared to other types of research such as drug trials where researchers and sponsors may favour a positive drug efficacy. That is, we are not aware of a mechanism by which the mean TOT would affect the likelihood of publication. The scatter plot of sample size vs. mean TOT (Fig. 6) showed no characteristic funnel shape, likely because the observed spread in mean TOT reflects study heterogeneity rather than imprecision of the mean TOT values.

Regarding the within-study analysis, where we assessed differences in mean TOT (D), publication bias is possible but not evident from our findings. For example, innovative types of TORs where publication bias may be expected (e.g., directional and peripheral TORs) showed near-zero effects (Fig. 3), thus indicating that small (null) effects were published. The number of studies per category of the within-subject analysis was too small to create funnel plots or perform a formal test of publication bias.

4.3. Conclusion and recommendations for future research

The meta-analysis included 129 studies that measured driver TOTs when resuming manual control after automated driving and investigated the effect of multiple factors related to the driver, the automation system, the human–machine interface, the driving situation, and the experimental setup. Notable findings are that the available time, a lack of experience with TORs, and using a handheld device were associated with substantially increased mean TOT. Although providing a take-over request yields a lower mean TOT than no take-over request at all (or a visual-only take-over request), the modality of the TORs had relatively minor effects on the mean TOT.

These findings have important implications for future research and design. In particular, instead of designing new types of take-over requests that may have only incremental effects on mean TOT, efforts could be made towards ensuring that drivers are prepared and trained to take over. Also, drivers should not be permitted to engage in handheld non-driving tasks if take-over situations can be urgent. Conducting non-driving tasks on a mounted (head-up) display could be a safer option in such cases.

Finally, our meta-analysis suggests that achieving a low mean TOT should not necessarily be a design target. We showed that drivers take more time (i.e., the mean TOT is higher) when they have more time (i.e., when the urgency is lower). Future engineering efforts should be directed towards ensuring that drivers actually have sufficient time, which could be done by building better sensors with larger look-ahead time or by using vehicle-to-vehicle communication.

5. Supplementary material

Supplementary materials are accessible via this link: https://doi.org/10.4121/uuid:75c28abe-6559-4273-85f4-927e969c1c59.

Acknowledgement

The authors were involved in the Marie Curie Initial Training Network (ITN) project HFAuto – Human Factors of Automated Driving (PITN-GA-2013-605817).

Appendix A

PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	1–3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	3
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4

Appendix A (*continued*)

Section/topic	#	Checklist item	Reported on page #
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	3
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	3
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4–6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	5
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	5
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, *meta*-regression), if done, indicating which were pre-specified.	5 Between-study analysis and mixed-effects model are *meta*-regressions
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	4
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	4
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Supplementary file
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	7–8
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, *meta*-regression [see Item 16]).	8–13

(*continued on next page*)

Appendix A (*continued*)

Section/topic	#	Checklist item	Reported on page #
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11, 13–14
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14–15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15, 16
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	16

References

References marked with an asterisk indicate studies included in the meta-analysis.

Bazilinskyy, P., & De Winter, J. C. F. (201). Auditory interfaces in automated driving: An international survey e13. *PeerJ Computer Science, 1*. https://doi.org/10.7717/peerj-cs.13.

*Befelein, D., Boschet, J., & Neukum, A. (2018). *Influence of non-driving-related tasks' motivational aspects and interruption effort on driver take-over performance in conditionally automated driving.* Paper presented at the 6th international conference on driver distraction and inattention, Gothenburg, Sweden. Retrieved from: http://ddi2018.org/wp-content/uploads/2018/10/S9.3-Befelein.pdf.

*Befelein, D., Naujoks, F., & Neukum, A. (2016). *How do naturalistic non-driving-related tasks influence take-over time and quality during highly automated driving.* Poster session presented at human factors and ergonomics society Europe chapter 2016 annual conference. Prague, Czech Republic.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*, 1088–1101. https://doi.org/10.2307/2533446.

*Belderbos, C.A.G., (2015). *Authority transition interface: A human machine interface for taking over control from a highly automated truck (MSc thesis).* The Netherlands: Delft University of Technology, Delft.

*Belwadi, A., Loeb, H., Shen, M., Shaikh, S., & McIntosh, C. W. (2018). Emergency autonomous to manual takeover in a driving simulator: teen vs. adult drivers–a pilot study (No. 2018-01-0499). SAE Technical Paper. https://doi.org/10.4271/2018-01-0499.

*Berghöfer, F., Purucker, C., Naujoks, F., Wiedemann, K., & Marberger, K. (2018). *Prediction of take-over time demand in highly automated driving. results of a naturalistic driving study.* Poster session presented at the human factors and ergonomics society Europe chapter 2018 annual conference.

*Biedermann, A., & Cieler, S. (2017). Bedeutung und Herausforderungen der Fahrerzustandserkennung im Kontext des vollautomatisierten Fahrens [Importance and challenges of driver condition detection in the context of fully automated driving]. In VDI-Tagung Der Fahrer im 21. Jahrhundert 2017, Der Mensch im Fokus technischer Innovationen (pp. 91–106). Braunschweig, DE: VDI-Gesellschaft Fahrzeug- und Verkehrstechnik.

*Blommer, M., Curry, R., Kochhar, D., Swaminathan, R., Talamonti, W., & Tijerina, L. (2015). The effects of a scheduled driver engagement strategy in automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59*, 1681–1685. https://doi.org/10.1177/1541931215591363.

*Blommer, M., Curry, R., Swaminathan, R., Tijerina, L., Talamonti, W., & Kochhar, D. (2017). Driver brake vs. steer response to sudden forward collision scenario in manual and automated driving modes. *Transportation Research Part F: Traffic Psychology and Behaviour, 45*, 93–101. https://doi.org/10.1016/j.trf.2016.11.006.

Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating composite scores: A literature review, application to content validity, and meta-analysis. *Organizational Research Methods, 10*, 689–709. https://doi.org/10.1177/1094428106294734.

Bond, C. F., Jr., Wiitala, W. L., & Richard, F. D. (2003). Meta-analysis of raw mean differences. *Psychological Methods, 8*, 406–418. https://doi.org/10.1037/1082-989X.8.4.406.

Bonett, D. G. (2008). Meta-analytic interval estimation for bivariate correlations. *Psychological Methods, 13*, 173–181. https://doi.org/10.1037/a0012868.

Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association, 71*, 791–799. https://doi.org/10.2307/2286841.

*Braunagel, C., Rosenstiel, W., & Kasneci, E. (2017). Ready for take-over? A new driver assistance system for an automated classification of driver take-over readiness. *IEEE Intelligent Transportation Systems Magazine, 9*, 10–22. https://doi.org/10.1109/MITS.2017.2743165.

*Bueno, M., Dogan, E., Selem, F. H., Monacelli, E., Boverie, S., & Guillaume, A. (2016). *How different mental workload levels affect the take-over control after automated driving.* Paper presented at 2016 IEEE 19th international conference on intelligent transportation systems (ITSC) (pp. 2040–2045), Rio de Janeiro, Brazil. https://doi.org/10.1109/ITSC.2016.7795886.

*Çapalar, J., & Olaverri-Monreal, C. (2017). *Hypovigilance in limited self-driving automation: Peripheral visual stimulus for a balanced level of automation and cognitive workload*. Paper presented at 2017 IEEE 20th international conference on intelligent transportation systems (ITSC) (pp. 27–31). IEEE.

*Carsten, O., & Jamson, A. H. (2011). Driving simulators as research tools in traffic psychology. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 87–96). Academic Press. https://doi.org/10.1016/B978-0-12-381984-0.10007-4.

*Clark, H., & Feng, J. (2017). Age differences in the takeover of vehicle control and engagement in non-driving-related activities in simulated driving with conditional automation. *Accident Analysis & Prevention, 106*, 468–479. https://doi.org/10.1016/j.aap.2016.08.027.

*Cohen-Lazry, G., Borowsky, A., & Oron-Gilad, T. (2017). The effects of continuous driving-related feedback on drivers' response to automation failures. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*, 1980–1984. https://doi.org/10.1177/1541931213601974.

*Cohen-Lazry, G., & Katzman, N. (2018). *Ipsilateral versus contralateral tactile alerts for take-over requests in highly-automated driving.* Paper presented at the 6th humanist conference, The Hague, Netherlands. Retrieved from http://www.humanist-vce.eu/fileadmin/contributeurs/humanist/TheHague2018/1-cohen-lazry.pdf.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30*, 116–127. https://doi.org/10.1037/h0076829.

*Cui, W., Zhou, R., Yan, Y., Ran, L., & Zhang, X. (2017). Effect of warning levels on drivers' decision- making with the self-driving vehicle system. In N. A. Stanton (Ed.), *Advances in human aspects of transportation* (pp. 720–729). Cham: Springer. https://doi.org/10.1007/978-3-319-60441-1_69.
*Damböck, D., Weissgerber, T., Kienle, M., & Bengler, K. (2013). *Requirements for cooperative vehicle guidance.* Paper presented at the 16th international IEEE conference on intelligent transportation systems-(ITSC) (pp. 1656–1661). The Hague, Netherlands. https://doi.org/10.1109/ITSC.2013.6728467.
Deeks, J. J., Macaskill, P., & Irwig, L. (2005). The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology, 58*, 882–893. https://doi.org/10.1016/j.jclinepi.2005.01.016.
De Winter, J. C. F., Van Leeuwen, P. M., & Happee, R. (2012). Advantages and disadvantages of driving simulators: A discussion. In A.J. Spink, F. Grieco, O.E. Krips, L.W.S. Loijens, L.P.J.J. Noldus, & P.H. Zimmerman (Eds.), Proceedings of measuring behavior (pp. 47–50). Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.388.1603&rep=rep1&type=pdf#page=47.
*De Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 196–217. https://doi.org/10.1016/j.trf.2014.06.016.
De Winter, J. C. F., Stanton, N. A., Price, J. S., & Mistry, H. (2016). The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. *International Journal of Vehicle Design, 70*, 297–324. https://doi.org/10.1504/IJVD.2016.076736.
De Winter, J. C. F., Zadpoor, A. A., & Dodou, D. (2014). The expansion of Google Scholar versus Web of Science: A longitudinal study. *Scientometrics, 98*, 1547–1565. https://doi.org/10.1007/s11192-013-1089-2.
*Dogan, E., Deborne, R., Delhomme, P., Kemeny, A., & Jonville, P. (2014). *Evaluating the shift of control between driver and vehicle at high automation at low speed: The role of anticipation.* Paper presented at the transport research arena (TRA) 5th conference: Transport solutions from research to deployment, Paris, France.
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629.
Einhorn, H. J., & Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance, 13*, 171–192. https://doi.org/10.1016/0030-5073(75)90044-6.
*Eriksson, A., Petermeijer, S. M., Zimmerman, M., De Winter, J. C. F., Bengler, K. J., & Stanton, N. (2019). Rolling out the red (and green) carpet: supporting driver decision making in automation-to-manual transitions. *IEEE Transactions on Human-Machine Systems, 49*, 20–31. https://doi.org/10.1109/THMS.2018.2883862.
*Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. *Human Factors, 59*, 689–705. https://doi.org/10.1177/0018720816685832.
*Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017). How the duration of automated driving influences take-over performance and gaze behavior. In C. M. Schlick, S. Duckwitz, F. Flemisch, M. Frenz, & S. Kuz (Eds.), *Advances in ergonomic design of systems, products and processes* (pp. 309–318). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-53305-5_22.
*Feldhütter, A., Hecht, T., Bengler, K. (2018). Fahrerspezifische Aspekte beim hochautomatisierten Fahren [driver-specific aspects in highly automated driving]. *FAT Schriftenreihe*.
*Feldhütter, A., Kroll, D., & Bengler, K. (2018). *Wake up and take over! The effect of fatigue on the take-over performance in conditionally automated driving.* Paper presented at the 21th international conference on intelligent transportation systems (ITSC) (pp. 2080–2085). https://doi.org/10.1109/ITSC.2018.8569545.
*Feldhütter, A., Segler, C., & Bengler, K. (2017). Does shifting between conditionally and partially automated driving lead to a loss of mode awareness? In N. A. Stanton (Ed.), *Advances in human aspects of transportation* (pp. 730–741). Cham: Springer. https://doi.org/10.1007/978-3-319-60441-1_70.
Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*, 381–391. https://doi.org/10.1037/h0055392.
Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., ... Santaguida, P. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the effective health care program. *Journal of Clinical Epidemiology, 64*, 1187–1197. https://doi.org/10.1016/j.jclinepi.2010.08.010.
*Funkhouser, K., & Drews, F. (2016). Putting the brakes on autonomous vehicle control: Responding to system breakdowns. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*, 1859–1863. https://doi.org/10.1177/1541931213601424.
Gold, C., & Bengler, K. (2014). *Taking over control from highly automated vehicles.* Paper presented at the 5th international conference on applied human factors and ergonomics (pp. 3662–3667).
*Gold, C., Berisha, I., & Bengler, K. (2015). Utilization of drivetime – Performing non-driving related tasks while driving highly automated. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59*, 1666–1670. https://doi.org/10.1177/1541931215591360.
*Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*, 1938–1942. https://doi.org/10.1177/1541931213571433.
*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*, 642–652. https://doi.org/10.1177/0018720816634226.
*Gold, C., Lorenz, L., & Bengler, K. (2014). *Influence of automated brake application on take-over situations in highly automated driving scenarios.* Paper presented at the FISITA 2014 world automotive congress, Maastricht, The Netherlands.
Gold, C., Lorenz, L., Damböck, D., & Bengler, K. (2013). Partially automated driving as a fallback level of high automation. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren, 6.
Gold, C., Happee, R., & Bengler, K. (2018). Modeling take-over performance in level 3 conditionally automated vehicles. *Accident Analysis & Prevention, 116*, 3–13. https://doi.org/10.1016/j.aap.2017.11.009.
Gold, C., Naujoks, F., Radlmayr, J., Bellem, H., & Jarosch, O. (2018). Testing scenarios for human factors research in level 3 automated vehicles. In N. A. Stanton (Ed.), *Advances in human aspects of transportation* (pp. 551–559). Cham: Springer. https://doi.org/10.1007/978-3-319-60441-1_54.
Green, M. (2000). "How long does it take to stop?" Methodological analysis of driver perception-brake times. *Transportation Human Factors, 2*, 195–216. https://doi.org/10.1207/STHF0203_1.
*Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). *Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving.* Paper presented at the 5th international conference on automotive user interfaces and interactive vehicular applications (pp. 210–217), Eindhoven, The Netherlands. https://doi.org/10.1145/2516540.2516554.
Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods, 3*, 486–504. https://doi.org/10.1037/1082-989X.3.4.486.
*Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. *Human Factors, 59*, 457–470. https://doi.org/10.1177/0018720816678714.
Higgins, J. P. T., & Green, S. (Eds.). (2005). Cochrane handbook for systematic reviews of interventions. The Cochrane collaboration.
*Homma, R., Wakasugi, T., & Kodaka, K. (2016). *Human-factors issues of transition from an automated driving system to manual driving.* Paper presented at the 23rd ITS world congress, Melbourne, Australia.
Horrey, W., & Wickens, C. (2007). In-vehicle glance duration: Distributions, tails, and model of crash risk. *Transportation Research Record: Journal of the Transportation Research Board, 2018*, 22–28. https://doi.org/10.3141/2018-04.
*Ito, T., Takata, A., & Oosawa, K. (2016). Time required for take-over from automated to manual driving (No. 2016-01-0158). SAE Technical Paper. https://doi.org/10.4271/2016-01-0158.
IBM Corporation (2016). IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

Jalali, S., & Wohlin, C. (2012). Systematic literature studies: Database searches vs. backward snowballing. In *Proceedings of the ACM-IEEE international symposium on empirical software engineering and measurement* (pp. 29–38). ACM. https://doi.org/10.1145/2372251.2372257.

*Jarosch, O., Kuhnt, M., Paradies, S., & Bengler, K. (2017). It's out of our hands now! Effects of non-driving related tasks during highly automated driving on drivers' fatigue. In *Human factors in driver assessment training and vehicle design* (pp. 319–325). Public Policy Center, University of Iowa. https://doi.org/10.17077/drivingassessment.1653.

*Josten, J., Zlocki, A., & Eckstein, L. (2016). Untersuchung der Bewältigungsleistung des Fahrers von kurzfristig auftretenden Wiederübernahmesituationen nach teilautomatischem, freihändigen Fahren [Investigation of drivers' handling performance in sudden take-over situations after semi-automated, hands-off driving]. FAT Schriftenreihe 289, Juli 2016.

Kaptein, N., Theeuwes, J., & Van der Horst, R. (1996). Driving simulator validity: Some considerations. *Transportation Research Record: Journal of the Transportation Research Board, 1550*, 30–36. https://doi.org/10.3141/1550-05.

*Kerschbaum, P., Lorenz, L., & Bengler, K. (2015). *A transforming steering wheel for highly automated cars.* Paper presented at the proceedings of the 2015 IEEE intelligent vehicles symposium (IV), Seoul, South Korea (pp. 1287–1292). https://doi.org/10.1109/IVS.2015.7225893.

*Kerschbaum, P., Omozik, K., Wagner, P., Levin, S., Hermsdörfer, J., & Bengler, K. (2017). How does a symmetrical steering wheel transformation influence the take-over process? In D. de Waard, A. Toffetti, R. Wiczorek, A. Sonderegger, S. Röttger, P. Bouchner, T. Franke, S. Fairclough, M. Noordzij, & K. Brookhuis (Eds.) Proceedings of the human factors and ergonomics society Europe chapter 2016 annual conference, Prague, Czech Republic (pp. 96–106).

Kim, J., Kim, H. S., Kim, W., & Yoon, D. (2018). *Take-over performance analysis depending on the drivers'* non-driving secondary tasks in automated vehicles. Paper presented at 2018 international conference on information and communication technology convergence (ICTC) (pp. 1364–1366). https://doi.org/10.1109/ICTC.2018.8539431.

*Kim, H. J., & Yang, J. H. (2017). Takeover requests in simulated partially autonomous vehicles considering human factors. *IEEE Transactions on Human-Machine Systems, 47*, 735–740. https://doi.org/10.1109/THMS.2017.2674998.

*Ko, S. M., & Ji, Y. G. (2018). How we can measure the non-driving-task engagement in automated driving: Comparing flow experience and workload. *Applied Ergonomics, 67*, 237–245. https://doi.org/10.1016/j.apergo.2017.10.009.

*Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics, 66*, 18–31. https://doi.org/10.1016/j.apergo.2017.07.006.

*Körber, M., Gold, C., Lechner, D., & Bengler, K. (2016). The influence of age on the take-over of vehicle control in highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 39*, 19–32. https://doi.org/10.1016/j.trf.2016.03.002.

*Körber, M., Schneider, W., & Zimmermann, M. (2015). *Vigilance, boredom proneness and detection time of a malfunction in partially automated driving.* Paper presented at the 2015 international conference on collaboration technologies and systems (CTS), Atlanta, GA (pp. 70–76). https://doi.org/10.1109/CTS.2015.7210402.

*Kreuzmair, C., Gold, C., & Meyer, M.-L. (2017). *The influence of driver fatigue on take-over performance in highly automated vehicles.* Paper presented at the 25th international technical conference on the enhanced safety of vehicles (ESV), Detroit, MI.

*Kuehn, M., Vogelpohl, T., & Vollrath, M. (2017). *Takeover times in highly automated driving (level 3).* Paper presented at the 25th international technical conference on the enhanced safety of vehicles (ESV), Detroit, MI.

*Langlois, S., & Soualmi, B. (2016). *Augmented reality versus classical HUD to take over from automated driving: An aid to smooth reactions and to anticipate maneuvers.* Paper presented at the 2016 IEEE 19th international conference on in intelligent transportation systems (ITSC), Rio de Janeiro, Brazil (pp. 1571–1578). https://doi.org/10.1109/ITSC.2016.7795767.

*Lapoehn, S., Dziennus, M., Utesch, F., Kelsch, J., Schieben, A., Dotzauer, M., ... & Köster, F. (2016). Interaction design for nomadic devices in highly automated vehicles. In Weyers, B., & Dittmar, A. (Eds.), Mensch und Computer 2016 – Workshopband. Aachen: Gesellschaft für Informatik e.V.. https://doi.org/10.18420/muc2016-ws08-0006.

*Lau, C.P., Harbluk, J.L., Burns, P.C., El-Hage, Y., (2018). *The influence of interface design on driver behavior in automated driving*. Victoria, BC: CARSP; The Canadian Association of Road Safety Professionals.

*Li, S., Blythe, P., Guo, W., & Namdeo, A. (2018). Investigation of older driver's takeover performance in highly automated vehicles in adverse weather conditions. *IET Intelligent Transport Systems, 12*, 1157–1165. https://doi.org/10.1049/iet-its.2018.0104.

Liu, K., & Green, P. (2017). The conclusion of a driving study about warnings depends upon how response time is defined. Proceedings of the human factors and ergonomics society annual meeting, 61, pp. 1876–1880. https://doi.org/10.1177/1541931213601949.

*Lorenz, L., Kerschbaum, P., & Schumann, J. (2014). *Designing take over scenarios for automated driving How does augmented reality support the driver to get back into the loop?*. Paper Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58, 1681–1685. https://doi.org/10.1177/1541931214581351.

*Louw, T., Merat, N., & Jamson, H. (2015). Engaging with highly automated driving: to be or not to be in the loop? Proceedings of the eighth international driving symposium on human factors in driver assessment, training and vehicle design, Salt Lake City, UT (pp. 190–196). https://doi.org/10.17077/drivingassessment.1570.

*Lu, Z., Zhang, B., Feldhütter, A., Happee, R., Martens, H. M., & De Winter, J. C. F. (2019). Beyond mere take-over requests: The effects of monitoring requests on driver attention, take-over performance, and acceptance. *Transportation Research Part F: Traffic Psychology and Behaviour, 63*, 22–37. https://doi.org/10.1016/j.trf.2019.03.018.

Lu, Z., Happee, R., Cabrall, C. D. D., Kyriakidis, M., & De Winter, J. C. F. (2016). Human factors of transitions in automated driving: A general framework and literature survey. *Transportation Research Part F: Traffic Psychology and Behaviour, 43*, 183–198. https://doi.org/10.1016/j.trf.2016.10.007.

*Manawadu, U. E., Hayashi, H., Ema, T., Kawano, T., Kamezaki, M., & Sugano, S. (2018). Tactical-level input with multimodal feedback for unscheduled takeover situations in human-centered automated vehicles. Paper presented at 2018 IEEE/ASME international conference on advanced intelligent mechatronics (AIM) (pp. 634–639). https://doi.org/10.1109/AIM.2018.8452227.

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google scholar, web of science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics, 12*, 1160–1177. https://doi.org/10.1016/j.joi.2018.09.002.

*McNabb, J. C. (2017). Warning a distracted driver: smart phone applications, informative warnings and automated driving take-over requests (Doctoral dissertation, Arizona State University).

*Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., & Bauer, W. (2015). Take-over requests for automated driving. *Procedia Manufacturing, 3*, 2867–2873. https://doi.org/10.1016/j.promfg.2015.07.788.

*Miller, D., Sun, A., Johns, M., Ive, H., Sirkin, D., Aich, S., & Ju, W. (2015). Distraction becomes engagement in automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59*, 1676–1680. https://doi.org/10.1177/1541931215591362.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*, 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.

Molenberghs, G., Kenward, M. G., & Verbeke, G. (2009). Discussion of likelihood inference for models with unobservables: Another view. *Statistical Science, 24*, 273–279. https://doi.org/10.1214/09-STS277B.

*Molnar, L. J., Pradhan, A. K., Eby, D. W., Ryan, L. H., St. Louis, R. M., Zakrajsek, J., ... Zhang, L. (2017). *Age-related differences in driver behavior associated with automated vehicles and the transfer of control between automated and manual control: a simulator evaluation (Report No. UMTRI 2017-4).* Ann Arbor, MI: University of Michigan Transportation Research Institute. Retrieved from: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/137653/UMTRI-2017-4%20.pdf?sequence=3&isAllowed=y.

*Morgan, P. L., Alford, C., Williams, C., Parkhurst, G., & Pipe, T. (2017). Manual takeover and handover of a simulated fully autonomous vehicle within urban and extra-urban settings. In N. A. Stanton (Ed.), *Advances in human aspects of transportation* (pp. 760–771). Cham: Springer. https://doi.org/10.1007/978-3-319-60441-1_73.

*Mueller, A., Ogrizek, M., Bier, L., & Abendroth, B. (2018). *Design concept for a visual, vibrotactile and acoustic take-over request in a conditional automated vehicle during non-driving-related tasks.* Paper presented at the 6th international conference on driver distraction and inattention, Gothenburg, Sweden. Retrieved from: http://ddi2018.org/wp-content/uploads/2018/10/S2.3-Mueller.pdf.

Naujoks, F., Befelein, D., Wiedemann, K., & Neukum, A. (2017). A review of non-driving-related tasks used in studies on automated driving. In N. A. Stanton (Ed.), *Advances in human aspects of transportation*. https://doi.org/10.1007/978-3-319-60441-1_52.

*Naujoks, F., Höfling, S., Purucker, C., & Zeeb, K. (2018). From partial and high automation to manual driving: Relationship between non-driving related tasks, drowsiness and take-over performance. *Accident Analysis & Prevention, 121*, 28–42. https://doi.org/10.1016/j.aap.2018.08.018.

*Naujoks, F., Mai, C., & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. In N. Stanton, S. Landry, G. Di Bucchianico, & A. Vallicelli (Eds.), Advances in human aspects of transportation, 7 (Part I), 431–438. Retrieved from: https://www.researchgate.net/publication/264734018_The_Effect_of_Urgency_of_Take-Over_Requests_During_Highly_Automated_Driving_Under_Distraction_Conditions.

Olaverri-Monreal, C., Kumar, S., & Dìaz-Àlvarez, A. (2018). *Automated driving: Interactive automation control system to enhance situational awareness in conditional automation.* Paper presented at 2018 IEEE intelligent vehicles symposium (IV) (pp. 1698–1703). https://doi.org/10.1109/IVS.2018.8500367.

Osburn, H. G., & Callender, J. (1992). A note on the sampling variance of the mean uncorrected correlation in meta-analysis and validity generalization. *Journal of Applied Psychology, 77*, 115–122. https://doi.org/10.1037/0021-9010.77.2.115.

*Othersen, I. (2016). Vom Fahrer zum Denker und Teilzeitlenker: Einflussfaktoren und Gestaltungsmerkmale nutzerorientierter Interaktionskonzepte für die Überwachungsaufgabe des Fahrers im teilautomatisierten Modus [From driver to thinker and part-time driver: Influencing factors and design features of user-oriented interaction concepts for the driver's monitoring task in semi-automated mode] (Doctoral dissertation, Technische Universität Braunschweig, Braunschweig, Germany).

Othersen, I., Petermann-Stock, I., Schoemig, N., & Fuest, T. (2017). Methoden und Wechselwirkung Kognitive Übernahmefähigkeit nach einer pilotierten Fahrt [Methods and interaction cognitive adoption after a pilot drive]. *ATZelektronik, 12*, 26–31. https://doi.org/10.1007/s35658-017-0068-4.

*Payre, W., Cestac, J., Dang, N.-T., Vienne, F., & Delhomme, P. (2017). Impact of training and in-vehicle task performance on manual control recovery in an automated car. *Transportation Research Part F: Traffic Psychology and Behaviour, 46*, 216–227. https://doi.org/10.1016/j.trf.2017.02.001.

*Payre, W., Cestac, J., & Delhomme, P. (2016). Fully automated driving: Impact of trust and practice on manual control recovery. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 58*, 229–241. https://doi.org/10.1177/0018720815612319.

*Petermann-Stock, I., Hackenberg, L., Muhr, T., & Mergl, C. (2013). Wie lange braucht der Fahrer–eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt [How long does the driver need - an analysis of take-over times from various non-driving tasks during highly automated traffic jam driving]. 6. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren. TÜV SÜD Akademie GmbH.

*Petermeijer, S. M., Cieler, S., & De Winter, J. C. F. (2017). Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. *Accident Analysis & Prevention, 99*, 218–227. https://doi.org/10.1016/j.aap.2016.12.001.

*Petermeijer, S. M., Bazilinskyy, P., Bengler, K., & De Winter, J. C. F. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied Ergonomics, 62*, 204–215. https://doi.org/10.1016/j.apergo.2017.02.023.

*Petermeijer, S. M., Doubek, F., & De Winter, J. C. F. (2017). Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants. Proceedings of the IEEE international conference systems, man, and cybernetics, Banff, AB, Canada. https://doi.org/10.1109/SMC.2017.8122827.

*Petermeijer, S. M., De Winter, J. C. F., & Bengler, K. (2016). Vibrotactile displays: A survey with a view on highly automated driving. *IEEE Transactions on Intelligent Transportation Systems, 17*, 897–907. https://doi.org/10.1109/TITS.2015.2494873.

*Pfromm, M. (2016). Menschbezogener Umgang mit Systemstörungen bei teilautomatisierter Fahrzeugführung [Human-related handling of system errors in semi-automated vehicle guidance] (Doctoral dissertation, Technische Universität Darmstadt, Darmstadt, Germany).

*Politis, I., Brewster, S., & Pollick, F. (2015). *Language-based multimodal displays for the handover of control in autonomous cars.* Paper presented at the 7th international conference on automotive user interfaces and interactive vehicular applications, Nottingham, UK (pp. 3–10). https://doi.org/10.1145/2799250.2799262.

*Politis, I., Brewster, S., & Pollick, F. (2017). Using multimodal displays to signify critical handovers of control to distracted autonomous car drivers. *International Journal of Mobile Human Computer Interaction, 9*, 1–16. https://doi.org/10.4018/ijmhci.2017070101.

*Politis, I., Langdon, P., Adebayo, D., Bradley, M., Clarkson, P. J., Skrypchuk, L., ... Stanton, N. (2018). *An evaluation of inclusive dialogue-based interfaces for the takeover of control in autonomous cars.* Paper presented at the 23rd international conference on intelligent user interfaces, Tokyo, Japan (pp. 601–606). https://doi.org/10.1145/3172944.3172990.

Radlmayr, J., & Bengler, K. (2015). Literaturanalyse und Methodenauswahl zur Gestaltung von Systemen zum hochautomatisierten Fahren [Literature analysis and method selection for designing systems for highly automated driving]. *FAT Schriftenreihe*. https://doi.org/10.13140/RG.2.1.2222.4401.

*Radlmayr, J., Fischer, F. M., & Bengler, K. (2019). The influence of non-driving related tasks on driver availability in the context of conditionally automated driving. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th congress of the international ergonomics association (IEA 2018). Advances in intelligent systems and computing: Vol. 823*. https://doi.org/10.1007/978-3-319-96074-6_32.

*Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58*, 2063–2067. https://doi.org/10.1177/1541931214581434.

*Radlmayr, J., Weinbeer, V., Löber, C., Farid, M., & Bengler, K. (2017). How automation level and system reliability influence driver performance in a cut-in situation. In N. A. Stanton (Ed.), *Advances in human aspects of transportation* (pp. 684–694). Cham: Springer. https://doi.org/10.1007/978-3-319-60441-1_66.

Riener, A. (2010). *Simulating on-the-road behavior using a driving simulator.* Paper presented at the 2010 third international conference on advances in computer-human interactions, Saint Maarten, Netherlands Antilles (pp. 25–31). https://doi.org/10.1109/ACHI.2010.10.

Risto, M., & Martens, M. H. (2014). Driver headway choice: A comparison between driving simulator and real-road driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 25*, 1–9. https://doi.org/10.1016/j.trf.2014.05.001.

*Roche, F., & Brandenburg, S. (2018). *Should the urgency of auditory-tactile takeover requests match the criticality of takeover situations?* Paper presented at the 21st international conference on intelligent transportation systems (ITSC) (pp. 1035–1040). https://doi.org/10.1109/ITSC.2018.8569650.

Rohatgi, A. (2017). WebPlotDigitizer [Web based tool]. Retrieved from: https://automeris.io/WebPlotDigitizer/index.html.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. https://doi.org/10.1002/0470870168.

*Sadeghian Borojeni, S., Boll, S. C. J., Heuten, W., Bülthoff, H. H., & Chuang, L. (2018). *Feel the movement: Real motion influences responses to take-over requests in highly automated vehicles.* Paper presented at the 2018 CHI conference on human factors in computing systems, Montreal QC, Canada, 246. https://doi.org/10.1145/3173574.3173820.

*Sadeghian Borojeni, S., Chuang, L., Heuten, W., & Boll, S. (2016). *Assisting drivers with ambient take-over requests in highly automated driving.* Paper presented at the 8th international conference on automotive user interfaces and interactive vehicular applications, Ann Arbor, MI (pp. 237–244). https://doi.org/10.1145/3003715.3005409.

*Sadeghian Borojeni, S., Weber, L., Heuten, W., & Boll, S. (2018). *From reading to driving: priming mobile users for take-over situations in highly automated driving.* Paper presented at the 20th international conference on human-computer interaction with mobile devices and services, 14. ACM. https://doi.org/10.1145/3229434.3229464.

SAE International. (2016). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicle (J3016) [Standard]*. Warrendale, PA: SAE International.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*, 507–514. https://doi.org/10.1016/j.neurobiolaging.2008.09.023.

*Schartmüller, C., Riener, A., & Wintersberger, P. (2018). *Steer-By-WiFi: Lateral vehicle control for take-overs with nomadic devices*. Paper presented at the 10th international conference on automotive user interfaces and interactive vehicular applications (pp. 121–126). ACM. https://doi.org/10.1145/3239092.3265954.

*Schartmüller, C., Riener, A., Wintersberger, P., & Frison, A. K. (2018). *Workaholistic: On balancing typing-and handover-performance in automated driving.* Paper presented at the 20th international conference on human-computer interaction with mobile devices and services, 16. ACM. https://doi.org/1145/3229434.3229459.

*Schittenhelm, H., & Nöcker, G. (2017). *Übernahmequalität bei Level 3 Systemen [Drivers take-over reaction times and quality with Level 3 systems].* Poster session presented at the 8th Tagung Fahrerassistenz, Einführung hochautomatisiertes Fahren, München, Germany. Retrieved from: https://www.researchgate.net/publication/321905188_Ubernahmequalitat_bei_Level_3_Systemen_Drivers_take-over_reaction_times_and_quality_with_Level_3_Systems.

Schmidt, F., & Hunter, J. (2015). Methods of meta-analysis. 55 City Road, London: SAGE Publications, Ltd. https://doi.org/10.4135/9781483398105.

*Schmidt, J. (2018). Detektion der Reaktionsbereitschaft beim hochautomatisierten Fahren [Detection of reaction readiness in highly automated driving] (Doctoral dissertation, Technical University of Munich, Munich, Germany).

*Schmidt, J., Dreißig, M., Stolzmann, W., & Rötting, M. (2017). The influence of prolonged conditionally automated driving on the take-over ability of the driver. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*, 1974–1978. https://doi.org/10.1177/1541931213601972.

Shen, S. (2016). Quantifying drivers' responses to failures of semi-autonomous vehicle systems (Doctoral dissertation, Clemson University).

*Sportillo, D., Paljic, A., Boukhris, M., Fuchs, P., Ojeda, L., & Roussarie, V. (2017). *An immersive virtual reality system for semi-autonomous driving simulation: A comparison between realistic and 6-DoF controller-based interaction*. Paper presented at the 9th international conference on computer and automation engineering (pp. 6–10). ACM. https://doi.org/10.1145/3057039.3057079.

*Sportillo, D., Paljic, A., Boukhris, M., Fuchs, P., Ojeda, L., & Roussarie, V. (2017). An immersive virtual reality system for semi-autonomous driving simulation: A comparison between realistic and 6-DoF controller-based interaction. In *Paper presented at the 9th international conference on computer and automation engineering* (pp. 6–10). ACM. https://doi.org/10.1145/3057039.3057079.

*Sportillo, D., Paljic, A., & Ojeda, L. (2018). Get ready for automated driving using virtual reality. *Accident Analysis & Prevention, 118*, 102–113. https://doi.org/10.1016/j.aap.2018.06.003.

Summala, H. (2000). Brake reaction times and driver behavior analysis. *Transportation Human Factors, 2*, 217–226. https://doi.org/10.1207/STHF0203_2.

*Talamonti, W. J. (2017). Human-machine interface development for modifying driver lane change behavior in manual, automated, and shared control automated driving (Doctoral dissertation, Wayne State University). Retrieved from: https://digitalcommons.wayne.edu/cgi/viewcontent.cgi?article=2746&context=oa_dissertations.

*Telpaz, A., Rhindress, B., Zelman, I., & Tsimhoni, O. (2015). *Haptic seat for automated driving: preparing the driver to take control effectively*. Paper presented at the 7th international conference on automotive user interfaces and interactive vehicular applications, Nottingham, UK (pp. 23–30). https://doi.org/10.1145/2799250.2799267.

*Telpaz, A., Rhindress, B., Zelman, I., & Tsimhoni, O. (2017). Using a vibrotactile seat for facilitating the handover of control during automated driving. *International Journal of Mobile Human Computer Interaction, 9*, 17–33. https://doi.org/10.4018/ijmhci.2017070102.

*Toffetti, A., Wilschut, E., Martens, M., Schieben, A., Rambaldini, A., Merat, N., & Flemisch, F. (2009). CityMobil: Human factor issues regarding highly automated vehicles on eLane. *Transportation Research Record: Journal of the Transportation Research Board, 2110*, 1–8. https://doi.org/10.3141/2110-01.

*Van den Beukel, A. P., Van der Voort, M. C., & Eger, A. O. (2016). Supporting the changing driver's task: Exploration of interface designs for supervision and intervention in automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 43*, 279–301. https://doi.org/10.1016/j.trf.2016.09.009.

*Van der Meulen, H., Kun, A. L., & Janssen, C. P. (2016). *Switching back to manual driving: how does it compare to simply driving away after parking?* Paper presented at the 8th international conference on automotive user interfaces and interactive vehicular applications, Ann Arbor, MI (pp. 229–236). https://doi.org/10.1145/3003715.3005452.

*Vlakveld, W., Van Nes, N., de Bruin, J., Vissers, L., & Van der Kroft, M. (2018). Situation awareness increases when drivers have more time to take over the wheel in a Level 3 automated car: A simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour, 58*, 917–929. https://doi.org/10.1016/j.trf.2018.07.025.

*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. https://doi.org/10.1016/j.trf.2018.03.019.

Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis & Prevention, 126*, 70–84. https://doi.org/10.1016/j.aap.2018.03.013.

Vogelpohl, T., Vollrath, M., Kühn, M., Hummel, T., Gehlert, T., (2016) *Übergabe von hochautomatisiertem Fahren zu manueller Steuerung [Transition from highly automated driving to manual control]*. Berlin: GDV.

*Walch, M., Lange, K., Baumann, M., & Weber, M. (2017). Autonomous driving: Investigating the feasibility of car-driver handover assistance. *International Journal of Mobile Human Computer Interaction, 9*, 58–74. https://doi.org/10.4018/IJMHCI.2017040104.

Walch, M., Mühl, K., Kraus, J., Stoll, T., Baumann, M., & Weber, M. (2017). From car-driver-handovers to cooperative interfaces: Visions for driver–vehicle interaction in automated driving. In G. Meixner & C. Müller (Eds.), *Automotive user interfaces: Creating interactive experiences in the car* (pp. 273–294). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-49448-7_10.

*Wan, J., & Wu, C. (2018a). The effects of lead time of take-over request and nondriving tasks on taking-over control of automated vehicles. *IEEE Transactions on Human-Machine Systems, 48*, 582–591. https://doi.org/10.1109/THMS.2018.2844251.

*Wan, J., & Wu, C. (2018b). The effects of vibration patterns of take-over request and non-driving tasks on taking-over control of automated vehicles. *International Journal of Human-Computer Interaction, 34*, 987–998. https://doi.org/10.1080/10447318.2017.1404778.

*Wandtner, B. (2018). Non-driving related tasks in highly automated driving-Effects of task characteristics and drivers' self-regulation on take-over performance (Doctoral dissertation, University of Würzburg, Würzburg, Germany). Retrieved from: https://opus.bibliothek.uni-wuerzburg.de/opus4-wuerzburg/frontdoor/deliver/index/docId/17395/file/Wandtner_Bernhard_AutomatedDriving.pdf.

*Wandtner, B., Schmidt, G., Schömig, N., & Kunde, W. (2018). *Non-driving related tasks in highly automated driving –Effects of task modalities and cognitive workload on take-over performance.* Paper presented at the AmE 2018 – automotive meets electronics; 9th GMM-symposium, Dortmund, Germany (pp. 8–13).

*Wandtner, B., Schömig, N., & Schmidt, G. (2018a). Effects of non-driving related task modalities on takeover performance in highly automated driving. *Human Factors, 60*, 870–881. https://doi.org/10.1177/0018720818768199.

*Wandtner, B., Schömig, N., & Schmidt, G. (2018b). Secondary task engagement and disengagement in the context of highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 58*, 253–263. https://doi.org/10.1016/j.trf.2018.06.001.

Warshawsky-Livne, L., & Shinar, D. (2002). Effects of uncertainty, transmission type, driver age and gender on brake reaction and movement time. *Journal of Safety Research, 33*, 117–128. https://doi.org/10.1016/S0022-4375(02)00006-3.

*Weinbeer, V., Baur, C., Radlmayr, J., & Bill, J., Muhr, T., & Bengler, K., (2017). *Highly automated driving: How to get the driver drowsy and how does drowsiness influence various take-over-aspects?* Paper presented at the 8th Tagung Fahrerassistenz, Einführung hochautomatisiertes Fahren, München, Germany.

*Wiedemann, K., Naujoks, F., Wörle, J., Kenntner-Mabiala, R., Kaussner, Y., & Neukum, A. (2018). Effect of different alcohol levels on take-over performance in conditionally automated driving. *Accident Analysis & Prevention, 115*, 89–97. https://doi.org/10.1016/j.aap.2018.03.001.

*Wintersberger, P., Riener, A., Schartmüller, C., Frison, A. K., & Weigl, K. (2018). *Let Me finish before I take over: Towards attention aware device integration in highly automated vehicles*. Paper presented at the 10th international conference on automotive user interfaces and interactive vehicular applications (pp. 53–65). ACM. https://doi.org/10.1145/3239060.3239085.

*Yang, Y., Gerlicher, M., & Bengler, K. (2018). How does relaxing posture influence take-over performance in an automated vehicle? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62*, 696–700. https://doi.org/10.1177/1541931218621157.

*Yang, Y., Götze, M., Laqua, A., Dominioni, G. C., Kawabe, K., & Bengler, K. (2017). A method to improve driver's situation awareness in automated driving. Paper presented at the human factors and ergonomics society Europe chapter 2017 annual conference (pp. 29–47). Retrieved from: https://www.hfes-europe.org/wp-content/uploads/2017/10/Yang2017.pdf.

*Yang, Y., Karakaya, B., Dominioni, G. C., Kawabe, K., & Bengler, K. (2018). An HMI concept to improve driver's visual behavior and situation awareness in automated vehicle. Paper presented at the 21st international conference on intelligent transportation systems (ITSC) (pp. 650–655). https://doi.org/10.1109/ITSC.2018.8569986.

*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. https://doi.org/10.1016/j.aap.2018.11.018.

Young, M. S., & Stanton, N. A. (2000). Journey's end: Will vehicle automation make skilled drivers redundant? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 44*, 89–92. https://doi.org/10.1177/154193120004400124.

*Young, M. S., & Stanton, N. A. (2007). Back to the future: Brake reaction times for manual and automated vehicles. *Ergonomics, 50*, 46–58. https://doi.org/10.1080/00140130600980789.

*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. https://doi.org/10.1016/j.aap.2015.02.023.

*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. https://doi.org/10.1016/j.aap.2016.04.002.

*Zeeb, K., Härtel, M., Buchner, A., & Schrauf, M. (2017). Why is steering not the same as braking? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 50*, 65–79. https://doi.org/10.1016/j.trf.2017.07.008.

*Zhang, B., Wilschut, E. S., Willemsen, D. M. C., Alkim, T., & Martens, M. H. (2017). The effect of see-through truck on driver monitoring patterns and responses to critical events in truck platooning. In N. A. Stanton (Ed.), *Advances in human aspects of transportation* (pp. 842–852). Cham: Springer. https://doi.org/10.1007/978-3-319-60441-1_81.

*Zhang, B., Wilschut, E. S., Willemsen, D. M. C., & Martens, M. H. (in press). Transitions to manual control from highly automated driving in non-critical truck platooning scenarios. *Transportation Research Part F: Traffic Psychology and Behaviour,* https://doi.org/10.1016/j.trf.2019.04.006.

Zuur, A.F., Leno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M. (2009). *Mixed effects models and extensions in ecology with R*. New York: NY: Spring Science and Business Media.