ORIGINAL ARTICLE

On credibility improvements for automotive navigation systems

**Florian Schaub · Markus Hipp ·
Frank Kargl · Michael Weber**

© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Automotive navigation systems are becoming ubiquitous as driver assistance systems. Vendors continuously aim to enhance route guidance by adding new features to their systems. However, we found in an analysis of current navigation systems that many share interaction weaknesses, which can damage the system's credibility. Such issues are most prevalent when selecting a route, deviating from the route intentionally, or when systems react to dynamic traffic warnings. In this work, we analyze the impact on credibility and propose improved interaction mechanisms to enhance perceived credibility of navigation systems. We improve route selection and the integration of dynamic traffic warnings by optimizing route comparability with relevance-based information display. Further, we show how bidirectional communication between driver and device can be enhanced to achieve a better mapping between device behavior and driver intention. We evaluated the proposed mechanisms in a comparative user study and present results that confirm positive effects on perceived credibility.

1 Introduction

Automotive navigation systems (ANS) have matured into a mainstream technology. While integrated ANS are mostly found in middle- and higher-range cars, cheaper portable navigation devices (PNDs) enable the addition of ANS into any vehicle. A navigation system's purpose is to support drivers in traveling from location A to destination B with route guidance. Modern ANS not only visualize the routing process on maps but contain additional features, like text-to-speech or advanced lane guidance with 3D visualization. Many devices can also receive up-to-date traffic information via the FM broadcast based TMC (Traffic Message Channel) and similar services.

ANS can be seen as support systems for safety critical situations, that is, the driving context. System errors or confusing commands can have significant consequences in cases where drivers rely blindly on the ANS. Especially in unfamiliar environments, drivers place higher confidence in navigation commands, while their self-confidence decreases [12]. In such situations, *gullibility errors* may occur, that is, the driver acts on an erroneous command perceived as credible. If drivers experience erroneous or misleading commands, they will trust ANS commands less in the future. The credibility of the ANS is damaged as a result, even due to small errors [8]. Credibility is a perceived quality that reflects the trustworthiness and expertise of a system. A loss of credibility in turn leads to

F. Schaub (✉) · M. Weber
Institute of Media Informatics, Ulm University,
89069 Ulm, Germany
e-mail: florian.schaub@uni-ulm.de

M. Weber
e-mail: michael.weber@uni-ulm.de

M. Hipp
Institute of Databases and Information Systems,
Ulm University, 89069 Ulm, Germany
e-mail: markus.hipp@uni-ulm.de

F. Kargl
Institute of Distributed Systems, Ulm University,
89069 Ulm, Germany
e-mail: frank.kargl@uni-ulm.de

F. Kargl
DIES Research Group, University of Twente,
P.O. Box 217, 7500 AE Enschede, The Netherlands

dismissal by the user. Therefore, exhibiting a high level of credibility is important to ensure continuous use of the system. High credibility ensures continuous benefit to the driver, but is also economically relevant for the ANS manufacturer, because low credibility may affect product or brand reputation [8].

Credibility issues not only occur in unfamiliar but also in familiar environments. In a familiar environment, a driver may form her own belief of the best route in a given situation. If the ANS does not support the driver's intention and does not make route recommendations sufficiently comprehensible, credibility of the ANS will also suffer. The driver may reject correct recommendations of the ANS, known as *incredulity errors* [8].

Navigation systems produce different kinds of erroneous messages that can impact credibility negatively. Many issues can be traced back to weaknesses in usability and interaction design. In this work, we provide an analysis of interaction weaknesses in current ANS based on an experimental study with PNDs (cf. Sect. 3) and relate them to core issues impacting credibility (cf. Sect. 4). But first, we provide background knowledge on credibility (cf. Sect. 2). We further propose a credibility-enhanced interaction design for ANS (cf. Sect. 5) focused on common task scenarios. Our evaluation in a comparative user study (cf. Sect. 6) validates that our contributions improve ANS credibility. We conclude the paper with an outlook on future directions in this line of work (Sect. 7).

2 Background on credibility

Fogg and Tseng [8] define credibility as a perceived quality comprised of a system's trustworthiness and expertise. Trustworthiness captures the perceived truthfulness of a system. Expertise captures the system's perceived knowledge and capabilities. Note that credibility is mainly concerned with believability, in contrast to trust that focuses on dependability [8, 26]. Fogg and Tseng [8] distinguish four types of credibility. *Presumed credibility* based on general assumptions about the system, for example, an ANS should find a route from A to B. The perceived quality of a system's hardware and interface determines *surface credibility*. *Reputed credibility* stems from experience reports by others, while *experienced credibility* results from personal experience.

Fogg and Tseng [8] further distinguish four credibility aspects users focus on when assessing credibility. *Device credibility* relates to a system's physical aspects, *functional credibility* to its functionality. These aspects are determined by the casing and routing engine of an ANS. *Interface credibility* and *information credibility* relate to the interaction experience and to how believable

information given by the system is [18, 26]. We mainly focus on the enhancement of the latter two.

In general, systems can gain credibility with users when they provide accurate information and lose it if provided information or system behavior is perceived as erroneous. Especially, small errors can have a disproportionally large effect on perceived credibility [8, 12]. Similarly, negative experiences affect trust to a greater extend than positive ones [16]. System credibility can be improved by facilitating understanding of system decisions [8, 19] and exhibiting reliable performance [16]. Interface design also influences credibility [8, 16]. For example, higher credibility is perceived for esthetic websites [23], a factor that could also be utilized by ANS. Too much trust, on the other hand, leads to *overreliance*, that is, the user neglects critical assessment of system commands [20]. Situational awareness is reduced and, thus, gullibility errors are facilitated. Therefore, credibility improvements should strive for an appropriate level of trust that matches the system's capabilities [16].

Most research on credibility focuses on website credibility [7, 18, 22, 23], only few work addresses credibility of ANS. Kantowitz et al. [12] showed that unreliable traffic information degrades the credibility of navigation systems. Pauzié [21] mentions the "legibility and understandability of messages" as a factor to gain benefits from ANS usage. Ross and Burnett [24] point out that "trust issues" arise if ANS directly start routing without showing the destination or an overview map. In the field of automotive HMI, most work focuses on general usability and interaction design aspects of ANS under consideration of the driving context's special requirements [1, 9, 15]. Multimodality [13, 17, 25] and driver attention and distraction [2, 4, 11] are prominent research topics. Proposed concepts could also positively affect credibility, for example, integrating landmarks in navigation commands to establish consistency with human navigational strategies [3]. However, effects on credibility are often not specifically evaluated and are not focus of related studies.

3 Analysis of current navigation systems

We assessed interaction weaknesses and credibility issues of current ANS in an experimental study with PNDs [10]. We tested five PNDs from different manufactures in real driving scenarios, ranging from low-end to top-range models representative of the German PND market in late 2009. We chose PNDs to ensure comparability between devices by mounting them in parallel in one car. We devised different scenarios to study ANS behavior in normal operation and in situations where driver and device intentions diverge. The first two scenarios simulate driving

in unfamiliar environments, while the detour scenarios simulate familiar environments:

Highway. The driver follows navigation commands on a long stretch of highway, including a short break for refueling.

Inner city. The driver is guided through a city center to a previously selected destination, including search for parking.

Detour (city). The driver takes a detour through the city to stop for coffee, while the PNDs advise to take the highway directly.

Detour (rural). The driver intentionally leaves the route on the highway for spontaneous sightseeing via rural roads.

Dynamic traffic warning. The integration of dynamic traffic warning messages is tested by driving on routes with reported traffic obstructions.

All devices were programmed with the same destination, and it was verified that suggested routes matched. For each scenario, audiovisual navigation commands of the devices were recorded. The road situation was filmed to facilitate correlation of navigation commands and driving situation, later on. We categorized recorded commands along two dimensions: *message correctness* (correct/false) and *driver anticipation* (expected/unexpected). A command is defined as *correct* if it corresponds to the driver's intended route. For example, if the driver leaves the route intentionally and the ANS insists on turning around, these messages would be considered *false*. As soon as the ANS recalculate and switch over to an alternative route, which matches the driver's intended route, subsequent commands would be considered *correct* again. Likewise, a message is *expected* if it can be anticipated by the driver. It fits the current situation and aligns with the driver's behavior. Surprising messages that do not fit the current situation appear *unexpected*. Note that this does not mean that the driver *knows* the content of an expected message in advance.

With these categories, we identified interaction weaknesses that cause mismatches between driver intention and ANS behavior. *Correct/expected* messages are normal navigation commands, and they constituted the majority of observed messages in our study (56 %).[1] *Correct/unexpected* messages (1 %) are not anticipated by the driver, but give correct advise. *False/unexpected* messages occurred rarely (2 %), for example, an erroneous "turn around" command caused by GPS reception issues, while driving with high speed on the highway. Such commands can cause critical incidents in case of overreliance and potentially

high credibility loss. *False/expected* messages are small errors that occur more frequently, for example, when leaving the route. If acting intentionally, the driver can *expect* these messages to be *false*. But persistent and repetitive messages of this kind can become a nuisance and reduce perceived expertise and trustworthiness. In our study, we observed 40 % *false/expected* messages in the detour (city) scenario, where the PNDs switched over to the detour route once it became shorter, and 93 % in the detour (rural) scenario, where no alternative route toward the destination was available.

Based on the results, we identified three common situations with high *false/expected* rates that exhibited prevalent interaction weaknesses across tested devices [10]:

Route selection. When ANS propose a route, the criteria for the recommendation are often unclear. The driver receives insufficient information to validate system recommendations. As a result, a mismatch between the driver's cognitive model of the best route and the proposed one may occur. Diminished credibility is the consequence.

Dynamic traffic information. When an ANS receives information about traffic obstructions and proposes an alternative route, provided information is often insufficient to make informed decisions. Choices were restricted to accepting the alternative or staying on the original route. Skepticism against these choices most likely increases with higher familiarity of the environment. Thus, a credibility decrease can be expected especially in familiar environments.

Deviation from route. ANS do not recognize if a driver deviates from the route by mistake or intentionally, for example, due to preferring familiar roads [14]. Thus, most current ANS either try to direct the driver back to the original route or recalculate the route to match the current driving direction. The new route may match the driver's intention for the moment, but ultimately leads toward the destination regardless of current driver intention. Inconsistent routing behavior with superfluous messages may be the result and reduce system credibility.

While the study only assessed a small set of ANS, the fact that the identified issues persisted across all tested devices irrespective of price segment and manufacturer suggests that the outlined interaction weaknesses deserve closer attention.

4 Impacting credibility

Concerning credibility, the interaction weaknesses in the identified scenarios can be broken down into a few core issues. *Insufficient information* and *insufficient choice* are

---

[1] We report only cumulative results as behavior differences were marginal between devices. Please see [10] for detailed results for each scenario.

salient issues in the route selection and dynamic traffic information scenarios. In the route deviation scenario, the issue is neither the ANS trying to fulfill its routing goal nor the driver intentionally ignoring navigation commands. *Insufficient communication capabilities* are the problem. The driver has no proper means to convey dynamic intention changes to the ANS. In the following, we elaborate on these issues and their impact on credibility.

4.1 Insufficient information

Insufficient information reduces verifiability of system decisions, which directly affects information credibility and as how believable presented information is perceived. When a mismatch between the driver's model of the best route and the system's proposed route occurs, the information is insufficient to convince the driver of the validity of the system's recommendation. If the driver cannot comprehend the system's actions, credibility is reduced.

Studies on website credibility have shown that additional information can enhance credibility [7] and that perceived information credibility encourages users to follow provided advice [18]. We hypothesize that the same effect can be achieved for ANS by providing more information about routing decisions. But due to the driving context, information presentation must be unobtrusive. Cognitive load has been shown to increase if drivers need to decipher presented information [21]. Therefore, information presentation must be optimized for the current situational context. ANS should provide the highest possible amount of useful information as concise as possible (*high entropy, low bandwidth*). Therefore, we propose a details on demand approach. Only most relevant information should be directly presented to the driver, with additional information being available on demand. Current ANS already provide information deemed relevant and even selectively provide additional information, for example, traffic message details, but information presentation is not optimized to the situational context. A details on demand approach tailored to the context would support the driver's assessment of the situation. This would help to resolve mismatches between the driver's and system's route models and, thus, retain credibility.

Another issue to consider is incorrect information in the ANS, that is, information that is inconsistent with the physical road situation, for example, road signs for one-way streets or speed limits, which are not reflected in the ANS. Optimized information presentation can help drivers estimate the trustworthiness of system recommendations [19, 20] and calibrate their trust to an appropriate level. An appropriate level of trust mitigates overreliance [16] and should reduce the risk of gullibility errors from incorrect information. In addition, other support systems such as visual road sign detection can be employed for correction of the ANS knowledge base.

4.2 Insufficient choice

The issue of insufficient choice is related to verifiability. When explicit decisions are required, drivers need sufficient information to validate the system's recommendation and evaluate alternatives. For example, based on provided information, drivers must decide if they want to circumvent a traffic jam or not. If the system does not properly support this decision-making process, drivers cannot make an informed decision (assuming they have no additional information from other resources like radio traffic service). If provided information is perceived as insufficient, drivers may not believe that the recommended route is optimal and the system will lose trust as a consequence [16]. If insufficient choices are offered, drivers will feel unsupported. In both cases, the system's interface credibility suffers. As Fogg and Tseng [8] put it, "an interface is likely to be perceived as less credible when it contradicts user expectation or mental models."

Many current ANS do not support evaluation of alternatives well. At initial route selection, most ANS provide only one route without alternatives [10]. In that case, drivers can only influence route selection prior to route calculation by setting few parameters, like fastest or most economic route. Similarly, when reacting to dynamic traffic warnings, the driver is commonly confronted with one detour option, which she can accept or reject. While some ANS manufacturers started offering multiple route choices for initial route selection, its impact on credibility has not been studied. We hypothesize that offering multiple alternatives enhances credibility because drivers feel supported in the decision process and in control. Providing more choice also means there are more options to properly align the driver's mental model and the system model, which reduces the likelihood of mismatches between the two. The system should preinterpret route alternatives to obtain a relevancy-based ordering and provide an explicit recommendation for the best route. The combination of giving a clear recommendation and enabling comparison with alternatives will likely suggest expertise to drivers.

4.3 Insufficient communication capabilities

ANS lack sufficient bidirectional communication capabilities. With many current systems, drivers only act actively when initially selecting the destination. While driving, driver interaction is reduced to interface control, like adjusting the zoom level or volume. Some inputs may not be allowed at all while driving in order to minimize driver distraction. Only occasionally ANS request driver input,

for example, when showing a dynamic traffic warning. Thus, while the system can convey dynamic information to the driver, drivers have very limited ability to convey dynamic intention changes to the system while driving. Typical examples for dynamic intention changes would be an unplanned trip to the grocery store or taking a detour for sightseeing. Current ANS are unable to adapt dynamically to such short-term changes in driver intention. Thus, when the driver deviates from the route, either a route alternative can be calculated or the driver is repeatedly asked to turn around. In that case, the driver's only options are to (1) reprogram the route, (2) deactivate routing, or (3) switch off the system. The first option entails an onerous process, which could be dangerous while driving. The second option often requires multiple steps and may only be performed if sufficiently annoyed. The third option is quick and effective, but the driver loses the moving map functionality.

Due to the lack of bidirectional communication, a mismatch between driver intention and system behavior ensues. Both perceived expertise and trustworthiness are likely damaged as a result, affecting functional and information credibility. By providing bidirectional communication capabilities during driving, drivers would be able to convey intention changes. Similar to how ANS prompt the driver for input in specific situations, ANS should also offer interaction capabilities to drivers in some situations, for example, when detecting a route deviation. Such interaction capabilities need to be unobtrusive and tailored to the driving context, so that drivers can make use of them if desired but are not forced to. By receiving explicit intention input from drivers, *false/expected* messages could be reduced, while retaining full functionality and utility. If utility is retained, drivers continue to trust the system because it performs as expected [16] and credibility remains intact.

5 Credibility enhancing interaction design

A holistic approach for interaction design is required to enhance ANS credibility. While improving credibility is the goal, applied concepts must not substantially increase driver distraction and cognitive load [1, 9]. Navigation is, and must remain, a secondary task in the driving context. Thus, credibility guidelines from other domains, such as website credibility [6, 7, 22, 23], cannot be directly applied to ANS. Following the discussion in the previous section, we aim to enhance credibility by

1. Providing information relevant to the given situation with additional information on demand to support verifiability of system decisions,

2. Offering alternatives to involve the driver in decision processes and to facilitate verifiability of system recommendations, and
3. Improving bidirectional communication to either let drivers convey intention changes when necessary or engage them in interaction when intentions are unclear.

In the following, we develop corresponding mechanisms for the three scenarios previously identified: route selection, dynamic traffic information, and deviation from the route. We will show that perceived credibility can be enhanced by addressing the interaction weaknesses in these scenarios with consistent improvements. Note that while focusing on these scenarios in this work, we designed the mechanisms with a broader task range in mind and are convinced that they are applicable beyond the task scenarios covered here. We involved a small number of drivers in the design process. They repeatedly provided feedback on drafts and variants of specific mechanisms, especially on the presentation of relevant information. Suggestions were used as initial guidance for designing interfaces with enhanced credibility.

5.1 Route selection

To enhance ANS credibility in the route selection process, we offer three route alternatives to the driver, along with information to facilitate comparison of these choices and validation of the system's recommendation. The driver must explicitly select one of the routes to start routing. This way, the driver should feel more involved in the route selection, which supports trustworthiness [24]. Also, the chance of routing errors due to wrongly selected destinations is reduced, which would negatively impact experienced credibility. We further encourage a validation of the entered destination by labeling the routes with names of characterizing streets as associative cues.

Figure 1a shows the route selection interface. The leftmost route is the system's recommendation. The route ordering provides a preference ranking that is additionally reflected by color coding [27]. Salient route characteristics are displayed to facilitate comparison. Route length, duration, and traffic density enable micro-level route comparison. Absolute and relative values are combined for time and length to keep information presentation concise and provide easily discernible tendencies. Traffic density is indicated with a traffic light metaphor. If required, more details can be accessed on demand with the respective *info* button for each route. The star rating summarizes the system's recommendation by mapping the system's internal ranking of each route to a star level. The star rating supports macro level comparison. Within a glance, drivers

Fig. 1 Improved user interface for route selection. **a** Multiple routes on main screen, **b** multiple routes on map screen

Fig. 2 Improved user interface for integrating dynamic traffic information. **a** Route with obstruction and alternatives, **b** details about traffic obstruction

can discern if the system rates all alternative routes similar or if the recommended route is rated superior. The different levels ease comparison while keeping cognitive load low. This supports the driver's understanding of the system's decisions and conveys expertise on multiple levels. Even if drivers reject the recommended route and choose an alternative, they should feel supported in their decision-making process.

In addition, a map overview of all routes is available via the *Map* button (cf. Fig. 1b). Coloring of routes is consistent with the main screen. The map encourages validation of the selected destination. The map centers on the current position [5], but zooming allows to assess the complete routes. A subset of route characteristics also enables comparison. To enlarge map space, bar charts combining absolute and relative values are used to compare route duration and length. Information in the bar charts is ordered in consistency with the main screen. The map view facilitates spatial comparison, while the main screen is optimized for multi-parameter comparison. By offering choice also in terms of comparison views, different driver preferences are supported.

5.2 Dynamic traffic information

The integration of dynamic traffic information is related to route selection, but interaction occurs while driving. Drivers may be alerted with an audio cue when the screen appears. Drivers should be provided with meaningful and verifiable choices when a traffic obstruction is reported. If drivers are well supported in the decision-making process,

provided information will suggest expertise and trustworthiness, while the decision itself (accepting or rejecting a detour) should not affect the system's credibility.

Consistent with the route selection scenario, three choices are offered when a relevant traffic warning is received: continue on the current route or select one of the two route alternatives. Available information is distributed between two screens to keep initial information presentation concise. The map view (cf. Fig. 2a) is the main screen because spatial information is most useful to quickly evaluate the extend of the obstruction and available detours. Detailed information about the traffic obstruction (cf. Fig. 2b) can be obtained via the *more info* button.

On the map, an icon indicates the obstruction and the traffic jam is highlighted (*purple*). The rating of the choices is conveyed with star ratings, ordering, and color coding, analogous to the previous scenario. Bar charts are also used here to facilitate comparison of length and duration. From the bar chart, the potential time saving of alternative routes is easily discernible. The reuse of familiar elements is expected to keep cognitive load relatively low even if the driver encounters dynamic traffic warnings rarely.

The details screen (cf. Fig. 2b) provides additional information about the traffic obstruction if the driver wants to validate the system's recommendation. Its cause and the message's actuality are shown. The time estimate for continuing is broken down into driving and waiting time to facilitate understanding of the estimate's nature. The development of the traffic jam is visualized by a small graph to give an intuition if it is increasing or decreasing. The details screen is a map overlay, so that route selection

Fig. 3 Improved user interface for deviations from route. **a** Pause routing dialog, **b** non-modal pause dialog
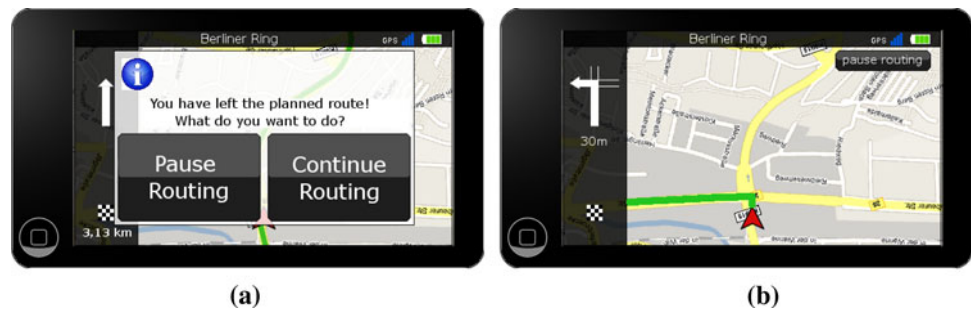
buttons remain available. Thus, the driver can assess the details and directly act upon them, which saves time and should positively affect functional credibility.

5.3 Deviation from route

Deviations from the route can be unintended or intended by the driver, but current ANS always assume an unintended deviation and cannot handle intended behavior. The result is potentially annoying turn around messages. However, supportive routing is essential in the case of unintended deviations. We propose an interaction design that supports both unintentional and intentional deviations by enabling bidirectional communication.

When a deviation is detected, our system informs drivers that they left the route with one concise voice command. The interface in Fig. 3a enables the driver to explicitly convey whether the deviation was intentional or unintentional by continuing or pausing navigation. Unless the driver reacts, the system assumes that the deviation was unintentional and continues routing. Thus, unintentionally acting drivers are not impaired by the dialog. After 15 s, the popup disappears to restore map visibility. The driver's inactivity is interpreted as implicit input, which is mapped to an unintentional deviation. However, the driver still has the option to pause routing with a button on the map (cf. Fig. 3b). If navigation is paused in either dialog, the ANS switch to free drive mode, that is, the current position is shown on the map without further routing commands. The button in Fig. 3b changes to *continue* and routing can be resumed anytime. The proposed interaction flow allows the driver to confirm intentional deviations while still fully supporting the driver in case of unintentional deviations, with the result that *false/expected* messages are eliminated and functional credibility is enhanced.

6 Evaluation

To evaluate the impact on ANS credibility, we implemented the developed concepts in a prototype system and conducted a comparative user study.

6.1 Setup and scenario description

An *experimental group* (*EG*) tested our prototype as the *experimental system* (*ES*), while a *control group* (*CG*) used a *control system* (*CS*). CS was consistently modeled after interaction and interface concepts representative of PNDs used in our original analysis [10]. The graphical design of CS and ES was homogenized to eliminate potential biases.

To ensure intra- and intergroup comparability of results, we opted for a desk-based laboratory study. ES/CS behavior was predefined and consistent for all participants. This allowed us to eliminate biases potentially caused by divergent routing behavior and driving variations in driving simulations and solely focus on the assessment of ANS credibility. We synchronized ES and CS actions to a recorded video of an actual drive, which was shown to participants while interacting with the ANS in the three scenarios. The simulated trip was 6.5 km long, consisting of 4.7 km rural and 1.8 km urban roads. Because participants were not actively driving, a disconnected task was introduced to generate a basic level of cognitive load. Arrows appeared shortly above the video and participants had to press respective arrow keys within a 5-s time frame. After each completed scenario, the test was interrupted and participants were presented with a questionnaire.

The *route selection* scenario was performed before driving started. Participants were asked to start routing to a destination. The destination was already prefilled and participants had to press *plan route*. ES presented the route selection (cf. Sect. 5.1) while CS started navigation directly on the "fastest route." In both cases, the same driving video was shown because route alternatives diverged only later on.

Next, the evaluation of the *dynamic traffic information* scenario followed in the rural section of the trip. Participants were allowed to practice the disconnected task for a few minutes without ANS interaction. Then, an accident report appeared and participants had to decide whether to circumvent the obstruction or not. CS provided a map with one detour option and duration/length information only for the detour, while ES offered multiple choices and details on demand (cf. Sect. 5.2). For both systems, continuing on

the route was the fastest option. The immediate driving situation was not affected by the user choice as the accident was further ahead on the route.

In the urban section, the *deviation from route* scenario was evaluated. Participants were told that they forgot to buy something and should turn right to reach the grocery store. The simulated trip deviated accordingly from the ANS route. CS routed back to the original route with voice commands, while ES offered the *pause* function (cf. Sect. 5.3).

6.2 Metrics

Credibility and related concepts are perceived qualities, which we assessed with a questionnaire after each scenario. Participants were asked to directly rate perceived *credibility*. But because credibility is a rather intangible concept, we also employed related terms suggested by Fogg and Tseng [8] to indirectly assess credibility. We asked for the perceived *believability* and *reliability* to measure trustworthiness and functional credibility. Thus, it is expected that results for believability and reliability also reflect effects on credibility. We further asked participants to rate the perceived *ability to influence system decisions (influence)* to support the assessment of *credibility*. Further items addressed experienced mental workload (*mental_load*) and asked to rate the amount of provided information (*info_amount*). Items were formulated as assertions, and participants were asked to rate them on a 5-point Likert scale from *does not apply at all* (1) to *applies fully* (5). Items concerning the amount of cognitive load or information could be rated from *too less* (1) to *too much* (5). Questionnaires were identical for both groups, except for additional items for EG to assess bar charts and star rating elements only present in ES.

6.3 Participants

The study was conducted with 42 participants (25 female, 17 male), equally distributed between groups. Participants were recruited from the university's student body, mainly from the disciplines Computer Science, Psychology, and Biology. All participants were aged 18–30, with the majority 25 or younger. Participants were randomly assigned to EG or CG. The demographic items *age*, *gender*, *technical affinity*, and *ANS ownership* were used as control variables. The distribution of *technical affinity* differentiated significantly between groups (two-tailed *t* test) and was, therefore, applied as covariant in all variance tests to compensate for non-uniform distribution.

Table 1 Results for the route selection scenario

Characteristic	$\overline{CG}$	σ_{CG}	$\overline{EG}$	σ_{EG}	p
Credibility	4.38	.80	4.48	.75	.490
Believability*	4.10	.70	4.57	.51	.026
Reliability	4.00	.89	4.34	.74	.160
Mental_load	1.29	.56	1.57	.68	.162
Info_amount*	2.76	.54	3.29	.56	.003
Influence**	2.33	1.24	3.95	1.16	<.001

* = 5 % significance level; ** = 0.1 % significance level

6.4 Results for route selection scenario

Variance analysis results for route selection are summarized in Table 1. While the difference in directly perceived credibility is not significant, believability was rated significantly higher by EG ($p = .026$), which can be interpreted as an indicator for higher credibility. The ability to influence the system's decisions was perceived significantly higher in EG ($p \leq .001$). Mental load was low in both groups, but the amount of available information was rated significantly higher by EG ($p = .003$). Thus, it can be concluded that the information in ES was more relevant to drivers. This is further supported by results for ES' additional interface elements. The star rating was perceived helpful ($\overline{EG} = 4.09, \sigma = .94$) and comprehensible ($\overline{EG} = 4.52, \sigma = .75$) while creating only low cognitive load ($\overline{EG} = 1.48, \sigma = .75$). The bar charts were considered comprehensible for time ($\overline{EG} = 4.05, \sigma = .89$) and length comparison ($\overline{EG} = 4.15, \sigma = .88$). They also positively impact believability of system recommendations, with time charts having a higher impact ($\overline{EG} = 4.15, \sigma = .75$) than length charts ($\overline{EG} = 4.10, \sigma = .64$). We conclude that facilitating comparability likely improves ANS credibility, because system recommendations are easier to validate.

6.5 Results for dynamic traffic information scenario

Table 2 summarizes the results. Credibility was perceived higher by EG, and believability was significantly higher ($p = .017$). Concerning the influence of available choices on believability (*choice_bel*), no difference between offering one (CS) or two alternatives (ES) was found. But considering that believability and the information amount were rated significantly higher by EG, it can be concluded that information has higher influence on believability than choice alone. This is further supported by the observation that only 38.1 % of CG chose to stay on the route (faster choice), in contrast to 61.9 % of EG; 57.1 % of EG used the details view, and 83.3 % of those chose to continue. This shows that providing additional information on demand

Table 2 Results for the dynamic traffic information scenario

Characteristic	$\overline{CG}$	σ_{CG}	$\overline{EG}$	σ_{EG}	p
Credibility	3.43	1.47	4.19	.93	.055
Believability*	3.38	1.28	4.24	.77	.017
Reliability	3.48	1.25	4.14	.79	.450
Mental_load	2.71	1.10	2.71	1.19	.563
Info_amount*	2.67	1.06	3.52	.68	.017
Influence	3.52	1.47	4.19	1.21	.127
Choice_bel	3.95	1.12	3.71	1.52	.309

* = 5 % significance level; ** = 0.1 % significance level

leads to better informed decisions, which translate to less frustration and higher experienced credibility. At the same time, the details on demand approach did not increase perceived cognitive load.

The ES' traffic jam characteristic (cf. Fig. 2b) was found to increase believability ($\overline{EG} = 4.47, \sigma = .74$). It was rated highly comprehensible ($\overline{EG} = 4.73, \sigma = .79$), while creating only low cognitive load ($\overline{EG} = 1.67, \sigma = .98$). The star rating received results similar to the previous scenario. Results for bar charts are also comparable, but slightly below the results of the previous scenario.

6.6 Results for deviation from route scenario

Table 3 summarizes results. While results show no significant difference for credibility, reliability was significantly better in EG ($p = .023$). As expected, CG rated the ability to influence the system quite low and stated that issued voice commands increased cognitive load (*acoustic*). Both items were rated significantly better by EG. Thus, the pause function significantly enhances the perceived reliability and, therefore, credibility. Furthermore, the CG results underline the negative effect of *false/ expected* messages.

6.7 Combined scales

The results for perceived characteristics such as *credibility* and *believability* do not exhibit consistent significance across scenarios. This instability is likely caused by their subjective nature. We presumed the subjectivity issue and, therefore, measured not just credibility but also the named related concepts that enabled some inferences for credibility.

In order to analyze general effects of the measured characteristics independent of specific scenarios, we combined items from all scenarios that measure one characteristic into a combined scale. For example, all items on credibility from the three scenarios were combined in one scale. This way, combined scales for believability, reliability, and influence were formed. Table 4 gives results of

Table 3 Results for the deviation from route scenario

Characteristic	$\overline{CG}$	σ_{CG}	$\overline{EG}$	σ_{EG}	p
Credibility	4.10	1.09	4.57	.60	.101
Believability	3.71	1.19	4.29	.96	.078
Reliability*	3.62	1.28	4.33	.86	.023
Mental_load	2.81	1.03	2.19	1.29	.281
Info_amount	2.95	.92	3.05	.38	.474
Influence**	2.00	1.18	4.33	1.24	<.001
Acoustic*	3.38	1.56	2.00	1.18	.002

* = 5 % significance level; ** = 0.1 % significance level

Table 4 Combined scales

Characteristic	$\overline{CG}$	$\overline{EG}$	F	p	α
Credibility*	3.81	4.33	59.67	.029	.655
Believability**	3.73	4.37	68.08	<.001	.649
Reliability*	3.70	4.29	68.46	.010	.618
Influence**	2.62	4.16	24.01	<.001	.842

the variance analysis on combined scales. All measured characteristics have been estimated significantly higher by the EG in the combined scales. Credibility of the improved system was perceived consistently higher in all scenarios, the combined *credibility* scale confirms this at a significant level. Note, however, that values for Cronbach's α are slightly lower than typically expected ($\alpha < .7$), which indicates a not fully consistent scale. A reason could be that participants might have attributed slightly different notions to credibility and related concepts across scenarios.

7 Conclusions

In this paper, we proposed three major concepts for ANS interaction design: (1) providing choices when decisions are required, (2) providing relevant information that facilitates comparison of alternatives, and (3) enabling bidirectional communication to let drivers convey intention. As a result, system decisions are easier verifiable, drivers feel involved and supported in navigation-related decisions, and *false/expected* navigation commands can be reduced.

The user study validates the positive effect on credibility of these concepts. However, it also shows that credibility is difficult to measure reliably. Assessment of related terminology is necessary, as already suggested by Fogg and Tseng [8], but also affects result reliability. The laboratory study provided unified conditions across groups, which simplified comparative evaluation of the developed concepts. At the same time, it introduced certain limitations. The actual risk of bad decisions usually experienced while

driving is missing, but may further impact credibility. The disconnected task did create basic cognitive load, as all participants were highly engaged in completing the task properly, but measurements did not allow statements about actual driver distraction. Especially, the influence of additional information on driver distraction requires further attention. Currently, the proposed concepts have only been assessed for three scenarios, and more studies under real driving conditions are required with ANS that fully implement these concepts to obtain more reliable results on their overall effect on ANS credibility. Long-term driver studies could also provide insights on how credibility and cognitive load develop over time. Comparing the impact on credibility of navigation commands and human driving suggestions could also provide interesting results.

The proposed concepts mainly improve credibility by enhancing explicit interaction. In future work, we plan to investigate the effects of implicit interaction on ANS credibility. Driving itself can be considered an implicit input channel, which allows inference of driving habits and potentially intention. ANS already contain sensors to measure location, heading, and speed to inform the navigation process. These parameters could be monitored over time to infer driving patterns and context. Furthermore, in-vehicle sensors for breaks, engine management, indicators, or steering could enrich context information. Implicit input could enhance credibility by optimizing system adaptation to the current context and tailoring explicit interaction accordingly. As one benefit, explicit interaction could be shifted to moments of relatively low cognitive load. For example, when the driver indicates while waiting at a traffic light although the route continues straight, the system could inquire the driver's intentions even before a deviation occurs. As another benefit, implicit interaction could be used to provide personalized route recommendations based on prior behavior. Future work is required to assess the potential and limitations of implicit interaction and potential benefits for ANS credibility.

References

1. Amditis A, Andreone L, Pagle K, Markkula G, Deregibus E, Rue MR, Bellotti F, Engelsberg A, Brouwer R, Peters B, De Gloria A (2010) Towards the automotive HMI of the future: overview of the AIDE-integrated project results. Trans Intell Transp Syst 11:567–578. doi:10.1109/TITS.2010.2048751

2. Bach KM, Jaeger MG, Skov MB, Thomassen NG (2009) Interacting with in-vehicle systems: understanding, measuring, and evaluating attention. In: Proceedings of the 23rd British HCI group annual conference on people and computers (BCS-HCI '09), pp 453–462. British Computer Society, Swinton, UK. http://portal.acm.org/citation.cfm?id=1671011.1671070

3. Burnett G (2000) The requirement for landmarks in vehicle navigation systems. J Navig 53(03):499–510. doi:10.1017/S0373463300001028

4. Chang JC, Lien A, Lathrop B, Hees H (2009) Usability evaluation of a Volkswagen Group in-vehicle speech system. In: Proceedings of the 1st international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '09), pp 137–144, ACM. doi:10.1145/1620509.1620535

5. Chittaro L (2006) Visualizing information on mobile devices. Computer 39(3):40–45. doi:10.1109/MC.2006.109

6. Fogg BJ, Marshall J, Laraki O, Osipovich A, Varma C, Fang N, Paul J, Rangnekar A, Shon J, Swani P, Treinen M (2001) What makes web sites credible?: A report on a large quantitative study. In: CHI'01:SIGCHI conference on human factors in computing systems, pp 61–68, ACM. doi:10.1145/365024.365037

7. Fogg BJ, Soohoo C, Danielson DR, Marable L, Stanford J, Tauber ER (2003) How do users evaluate the credibility of web sites?: a study with over 2,500 participants. In: DUX '03: conference on designing for user experiences, pp 1–15, ACM, New York. doi:10.1145/997078.997097

8. Fogg BJ, Tseng H (1999) The elements of computer credibility. In: CHI '99: SIGCHI conference on human factors in computing systems, pp 80–87, ACM. doi:10.1145/302979.303001

9. Green P, Levison W, Paelke G, Serafin C (1993) Preliminary human factors design guidelines for driver information systems. Technical Report UMTRI-93-21, University of Michigan Transportation Research Institute (UMTRI)

10. Hipp M, Schaub F, Kargl F, Weber M (2010) Interaction weaknesses of personal navigation devices. In: 2nd international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '10), ACM. doi:10.1145/1969773.1969796

11. Jensen BS, Skov MB, Thiruravichandran N (2010) Studying driver attention and behaviour for three configurations of gps navigation in real traffic driving. In: Proceedings of the 28th international conference on human factors in computing systems (CHI '10), pp 1271–1280, ACM. doi:10.1145/1753326.1753517

12. Kantowitz B.H., Hanowski RJ, C. Kantowitz S (1997) Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. Hum Fact 39(2). doi:10.1518/001872097778543831

13. Kaufmann C, Risser R, Geven A, Sefelin R (2008) Effects of simultaneous multi-modal warnings and traffic information on drivers behaviour. In: HUMANIST Proceedings 2008

14. Khattak AJ, Schofer JL, Koppelman FS (1993) Commuters' enroute diversion and return decisions: analysis and implications for advanced traveler information systems. Transp Res Part A: Policy Pract 27(2):101–111. doi:10.1016/0965-8564(93)90066-T

15. Lavie T, Oron-Gilad T, Meyer J (2011) Aesthetics and usability of in-vehicle navigation displays. Int J Hum Comput Stud 69:80–99. doi:10.1016/j.ijhcs.2010.10.002

16. Lee JD, See KA (2004) Trust in automation: designing for appropriate reliance. Hum Fact 46(1):50–80. doi:10.1518/hfes.46.1.50_30392

17. Lee WC, Ma MC, Cheng BW (2010) Field comparison of driving performance using a portable navigation system. J Navig 63(01):39–50. doi:10.1017/S0373463309990221

18. McKnight DH, Kacmar CJ (2007) Factors and effects of information credibility. In: 9th international conference on electronic commerce (ICEC '07), pp 423–432, ACM. doi:10.1145/1282100.1282180

19. Muir BM (1987) Trust between humans and machines, and the design of decision aids. Int J Man Mach Stud 27(5-6):527–539. doi:10.1016/S0020-7373(87)80013-5

20. Parasuraman R, Riley V (1997) Humans and automation: use, misuse, disuse, abuse. Hum Fact 39(2):230–253. doi:10.1518/001872097778543886

21. Pauzie A (2008) A method to assess the driver mental workload: the driving activity load index (DALI). IET Intell Transp Syst 2(4):315–322. doi:10.1049/iet-its:20080023

22. Rieh SY, Danielson DR (2007) Credibility: A multidisciplinary framework. Annu Rev Inf Sci Technol (ARIST) 41:307–364

23. Robins D, Holmes J (2008) Aesthetics and credibility in web site design. Inf Process Manag 44(1):386–399. doi:10.1016/j.ipm.2007.02.003

24. Ross T, Burnett G (2001) Evaluating the human-machine interface to vehicle navigation systems as an example of ubiquitous computing. Int J Hum Comput Stud 55(4):661–674. doi:10.1006/ijhc.2001.0495

25. Rydström A, Grane C, Bengtsson P (2009) Driver behaviour during haptic and visual secondary tasks. In: 1st international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '09), pp 121–127, ACM. doi:10.1145/1620509.1620533

26. Tseng S, Fogg BJ (1999) Credibility and computing technology. Commun ACM 42(5):39–44. doi:10.1145/301353.301402

27. Ware C (2004) Information Visualization, 2nd edn. Morgan Kaufmann, Los Altos