Personal and Ubiquitous Computing manuscript No.
(will be inserted by the editor)

On Credibility Improvements for Automotive Navigation Systems

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

Received: date / Accepted: date

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 which confirm positive effects on perceived credibility.

Keywords Automotive navigation systems · ANS · credibility · automotive HMI · HCI · interaction design

Florian Schaub and Michael Weber
Institute of Media Informatics, Ulm University, 89069 Ulm, Germany
Tel.: +49-731-50-31301
Fax: +49-731-50-31302
E-mail: florian.schaub@uni-ulm.de, michael.weber@uni-ulm.de

Markus Hipp
Institute of Databases and Information Systems, Ulm University, 89069 Ulm, Germany
Tel.: +49-731-50-24131
Fax: +49-731-50-24134
E-mail: markus.hipp@uni-ulm.de

Frank Kargl
DIES Research Group, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
Tel.: +31-53-489-4302
Fax: +31-53-489-3793
E-mail: f.kargl@utwente.nl

in Personal and Ubiquitous Computing Theme Issue on Automotive User Interfaces and Interactive Applications in the Car, 2012 Springer
Author pre-print draft. The final publication is available at www.springerlink.com. http://dx.doi.org/10.1007/s00779-012-0519-0

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, i.e., 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, i.e., 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 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. Sec. 3) and relate them to core issues impacting credibility (cf. Sec. 4). But first we provide background knowledge on credibility (cf. Sec 2). We further propose a credibility enhanced interaction design for ANS (cf. Sec. 5–7) focused on common task scenarios. Our evaluation in a comparative user study (cf. Sec. 8) validates that our contributions improve ANS credibility. We conclude the paper with an outlook on future directions in this line of work (Sec. 9).

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 skill. Note that credibility is mainly concerned with believability, in contrast to trust which focuses on dependability [8,24]. Fogg and Tseng [8] distinguish four types of credibility. *Presumed credibility* based on general assumptions about the system, e.g., 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 [17,24]. 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]. System credibility can be improved by facilitating understanding of system decisions [8,18]. Interface design also influences credibility [8]. For example, higher credibility is perceived for aesthetic websites [21], a factor that could also be utilized by ANS.

Most research on credibility focuses on website credibility [7,17,20,21], only few work addresses credibility of ANS. Kantowitz et al. [12] showed that unreliable traffic information degrades the credibility of navigation systems. Pauzié [19] mentions the "legibility and understandability of messages" as a factor to gain benefits from ANS usage. Ross and Burnett [22] 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 [15,9,1]. For example multimodality [16,13,23] or driver attention and distraction [2,11,4] to name two prominent topics. Proposed concepts could also positively affect credibility, e.g., 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

In order to assess interaction weaknesses of current ANS, we performed an experimental study with PNDs [10]. We tested five PNDs, ranging from low-end to top-range models, in real driving scenarios. Devices were mounted in parallel in one car to ensure comparability of their commands and reactions. We chose the following driving scenarios to assess ANS behavior in normal operation and to study their reactions when driver and device intentions diverge. The first two scenarios simulate driving in unfamiliar environments, while the detour scenarios simulate familiarity with the environment.

Highway. The driver follows navigation commands on a long stretch of motorway, 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 a coffee, while the ANS advise to take the highway.

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.

We recorded audiovisual navigation commands and categorized them by *message correctness* (correct / false) and *driver anticipation* (expected / unexpected). A command is defined as *correct* if it corresponds to the driver's intended route. So the system insisting on turning back to the original route against driver intention would be considered *false*. When the ANS system switches over to an alternative route aligned again with driver intention, 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, but not necessarily with her intention. 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. In total, we observed 56% *correct / expected* and 41% *false / expected* messages. While only 1% were labeled *correct / unexpected* messages and 2% *false / unexpected*. For detailed results, we refer the interested reader to our study [10].

We used these categories to identify interaction weaknesses that cause mismatches between driver intention and ANS behavior. *Correct / expected* messages are desirable, they constitute the majority of observed messages. *False / unexpected* messages occur rarely but give erroneous commands which can lead to critical incidents and a potentially high loss of credibility. For example, an erroneous "turn around" command while driving on the highway. *False / expected* messages are small errors that commonly occur, e.g., when leaving the route. If acting intentionally, the driver can *expect* these messages to be *false*. But persistent and repetitive messages of that kind become a nuisance and reduce perceived expertise and trustworthiness. Based on the study results, we identified three common task scenarios with high *false / expected* rates across devices that exhibit prevalent interaction weaknesses [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 updated traffic information and proposes an alternative route, provided information is often insufficient to make informed decisions. Furthermore, choices are restricted to accepting the alternative or staying on the original route. Thus, drivers may be sceptic about provided choices. This scepticism 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 original route by mistake or intentionally, e.g., due to preferring certain familiar roads [14]. Current ANS try to direct the driver back to the original route until they switch over to an alternative, which again could or could not match the driver's intention. Often, erratic routing behavior with superfluous messages is the result. Inconsistent and unsupportive behavior reduces system credibility.

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 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 [17]. We hypothesise 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 [19]. 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, e.g., 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 and help resolve mismatches between the driver's and system's route models. Credibility would be maintained. In order to optimize information presentation accordingly, it is essential to know what information is relevant to a driver in a given situation. Section 6 addresses this issue.

4.2 Insufficient Choice

The issue of insufficient choice is also 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 the driver must decide if she wants to circumvent a traffic jam or not. If the system does not properly support the driver in this decision making process, she cannot make an informed decision (assuming she has 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. If insufficient choices are offered, drivers will feel unsupported. In both cases, the system's interface credibility suffers. As Fogg and Tseng put it "an interface is likely to be perceived as less credible when it contradicts user expectation or mental models" [8].

Current ANS do not support evaluation of alternatives well. At initial route selection, most ANS provide only one route without alternatives. 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. We hypothesize that offering multiple alternatives will enhance credibility because drivers will 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 pre-interpret 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. After that, driver interaction is reduced to interface control, like adjusting the zoom level or volume. Only occasionally ANS request driver input, e.g., 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. The driver is continuously directed back to the original route until the system switches over to a recalculated route. 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 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, e.g., when detecting a route deviation. 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 and credibility remains intact. At the same time, such interaction capabilities need to be unobtrusive, so that drivers can make use of them if desired but are not forced to.

5 Approach and Methodology

A holistic approach for interaction design is required to enhance ANS credibility. It is especially important to take special requirements for in-vehicle systems into account [1,9]. While improving credibility is the goal, applied concepts must not substantially increase the driver's cognitive load. Navigation is and must remain a secondary task in the driving context. Thus, credibility guidelines from other domains, such as website credibility [6,7, 21,20], cannot be directly applied to ANS. Following the discussion in the previous section, we aim to enhance credibility by

1. providing information relevant in a 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 let drivers convey intention changes when necessary or engage them in interaction when intentions are unclear.

We develop corresponding mechanisms for the three scenarios previously identified: route selection, dynamic traffic information, and deviation from route. We will show that by addressing the interaction weaknesses in these scenarios with consistent improvements perceived credibility can be enhanced. 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.

The methodology of our work is as follows. We first performed an exploratory prestudy to gain insights on the relevance of information items in different scenarios (cf. Sec. 6). Building on those results, we developed an improved interaction design for ANS in an user-centered design process (cf. Sec. 7) and assessed the effect on credibility of the proposed improvements in a comparative user study (cf. Sec. 8).

6 Exploratory Study on Information Relevance

ANS have an abundant range of information available that could be presented to drivers. For example, a route can be characterized by its *length* (absolute and relative to alternative routes), its *duration* (absolute and relative to alternative routes), its *road characteristics* (e.g., curvy road, tollway), the *road type ratio* (47% interstate, 37% highway, 16% urban), and *traffic density*.

Dynamic traffic information adds more items. Typically, the *location* and *type* of a traffic obstruction, the *current traffic jam length* and the *estimated length on arrival*, the current and estimated *waiting time*, the *estimated time saving* when bypassing the traffic jam, and the *actuality* of the traffic message (i.e., how current is the information).

Which information is relevant in a given situation depends on individual drivers. But a coarse relevance classification of information items should facilitate enhancements in information presentation to improve credibility.

In an online questionnaire, we asked participants to rate subjective relevance of information items on a 5-point Likert scale (1 = *not important*, 5 = *very important*). Questions focused on relevant route characteristics and relevant information about traffic obstructions to inform the interface design for the route selection and dynamic traffic information scenarios.

6.1 Results

The study was completed by 30 participants. Participants were mostly male (90%) and 30% own an ANS. Due to sample size ($n = 30$) the results are not representative but are considered useful as initial guidance for designing interfaces with enhanced credibility, as supported by results in Section 8.

6.1.1 Route characteristics

Route duration was rated the most relevant route characteristic ($\bar{x}$=4.4, σ=.9), closely followed by *time difference between routes* ($\bar{x}$=4.3, σ=.9). The *route length* ($\bar{x}$=4.2, σ=1.0) and *relative route length* ($\bar{x}$=4.1, σ=1.1) were rated similarly high. Thus, duration and length information are most relevant in characterising a route. Relative information is probably regarded relevant because it facilitates comparative evaluation of different routes. This is further underlined by the importance of *providing multiple route alternatives* in the route selection process ($\bar{x}$=3.6, σ=1.0). Hereby, the majority (53.3%) preferred the display of three routes (<3=23.3%, >3=3.3%, 20.3% no preference). *Traffic density* ($\bar{x}$=4.1, σ=.6) outweighs *road type ratio* ($\bar{x}$=3.3, σ=.8) and *road characteristics* ($\bar{x}$=2.7, σ=1.3).

6.1.2 Traffic obstruction characteristics

The *actuality* of traffic information was rated most relevant ($\bar{x}$=4.8, σ=.4). *Estimated time saving* of an alternative route was also rated high ($\bar{x}$=4.4, σ=.6), which conforms to results of previous studies [14]. Related are *estimated wait time* in the traffic jam ($\bar{x}$=4.2, σ=.9), *traffic jam length (current)* ($\bar{x}$=4.3, σ=.8) and estimated *traffic jam length on arrival* ($\bar{x}$=3.9, σ=1.0). The development of a traffic jam (*traffic jam characteristic*) was also rated high ($\bar{x}$=4.0, σ=.8). While current dynamic traffic services, like TMC, do not provide continuous updates for specific obstructions, future services, like TPEG, may support this.

For alternative routes, the *traffic density* ($\bar{x}$=3.9, σ=.8), the *length* ($\bar{x}$=3.6, σ=1.2), and the *relative length* compared to the current route ($\bar{x}$=3.5, σ=1.3) are relevant. *Road type ratio* ($\bar{x}$=2.7, σ=1.2) and *road characteristic* ($\bar{x}$=2.3, σ=1.2) are less important.

7 Credibility Enhancing Interaction Design

In the following, we propose credibility enhanced interaction design for the three task scenarios. We followed an iterative user-centered design approach, with a small number of drivers constantly providing feedback on early drafts and variants to refine concepts.

7.1 Route Selection

To enhance ANS credibility in the route selection process, we offer the driver three route alternatives along with information to facilitate comparison of these choices and validation of the system's recommendation. Rather than start routing automatically after a destination has been entered, the driver explicitly selects a route. This way, the driver should feel more involved in the route selection, which supports trustworthiness [22]. 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.

Fig. 1(a) shows the route selection interface. The route on the left is the system's recommendation. The route ordering reflects a ranking also reflected by color-coding [25]. Salient route characteristics are displayed to facilitate comparison. Based on the exploratory study results, we provide route length, duration, and traffic density for 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 could be accessed per route on demand with the respective *info* button. The star rating summarizes the system's recommendation to support macro level comparison, e.g., within a driver's quick glance. 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 the driver rejects the recommended route and chooses an alternative, she should feel supported in her decision making.

(a) Multiple routes on main screen.

(b) Multiple routes on map screen.

Fig. 1 Improved user interface for route selection.

In addition, a map overview of all routes is available via the *Map* button (cf. Fig. 1(b)). Routes are colored in consistency 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. 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.

7.2 Dynamic Traffic Information

The integration of dynamic traffic information is related to route selection. The driver should be provided with meaningful and verifiable choices. If the driver is well supported in her 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 dynamic traffic warning is received: continue on the current route or select one of two route alternatives. Available information is distributed between two screens to keep initial information presentation concise. The map view (cf. Fig. 2(a)) is the main screen because spatial information is most useful to evaluate the extend of the obstruction and available detours. Detailed information about the traffic obstruction (cf. Fig 2(b)) 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 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 is easily discernible, which was rated highly in our exploratory study. 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. 2(b)) provides information about the traffic obstruction to enable validation of 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

(a) Route with obstruction and alternatives. (b) Details about traffic obstruction.

Fig. 2 Improved user interface for integrating dynamic traffic information.

screen is an overlay on the map, so that the buttons to choose a route remain available. Thus, the driver can assess the details and directly act upon them, which saves time and should positively affect functional credibility.

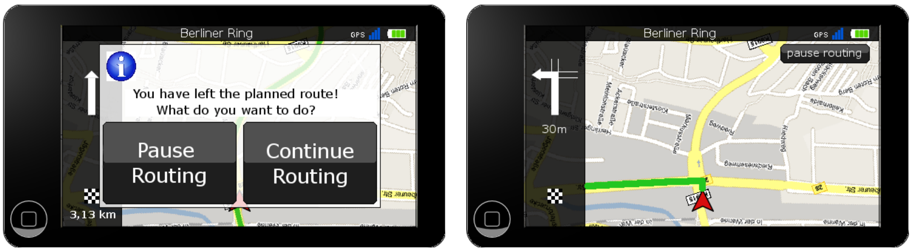
7.3 Deviation from Route

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 are potentially annoying turn around messages. However, supportive routing is essential in the case of unintended deviations. We propose an interaction design which supports both unintentional and intentional deviations by enabling bidirectional communication.

When a deviation is detected, our system informs the driver that she left the route with a concise voice command. The interface in Fig. 3(a) 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 towards the destination. Thus, unintentionally acting drivers are not impaired by the dialog. After 15s, 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 the routing with a button on the map (cf. Fig. 3(b)). If navigation is paused in either dialog, the ANS switches to free drive mode, i.e., the current position on the map is shown but no more routing commands are given. The button in Fig. 3(b) 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.

8 Evaluation

To evaluate the impact on ANS credibility, we implemented the concepts proposed in Section 7 in a prototype system. We conducted a comparative user study in which an *experimental group (EG)* tested our prototype as the *experimental system (ES)*, while a *control group (CG)* used a *control system (CS)*. The CS was consistently modelled after interaction and



(a) Pause routing dialog.

(b) Non-modal pause dialog.

Fig. 3 Improved user interface for deviations from route.

interface concepts of representative PNDs used in our original analysis [10]. The graphical design of CS and ES was homogenized to eliminate potential biases.

To ensure reliability as well as intra-group and inter-group comparability of results, we opted for a desk-based lab study. ES/CS behavior was predefined and consistent for all participants, with conditional branches to simulate interaction. This allowed us to eliminate biases potentially caused by divergent routing behavior and driving variations in actual 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 the participants while interacting with the ANS. The simulated trip was 6.5 km long, consisting of 4.7 km rural roads and 1.8 km urban roads. The *route selection* scenario was performed before driving started. Here, the CS calculated the “fastest route” and started navigation directly. The rural section included the evaluation of the *dynamic traffic information* scenario. Participants had to decide whether to circumvent an obstruction or not. Here, the CS provided a map with one detour option and duration/length information only for the detour. For both systems, continuing on the original route was the fastest option. In the urban section, the *deviation from route* scenario was evaluated. When the simulated trip deviated from the ANS’ route, CS routed back to the original route, while ES offered the *pause* function.

During the test, artificial cognitive load was created by a disconnected task to assess experienced cognitive load. Participants had to press specific arrow keys displayed shortly.

Participants were randomly assigned to EG or CG. An initial questionnaire asked for demographic information and technology affinity, before participants interacted with the systems. Interaction was recorded and each scenario was followed by a dedicated questionnaire. Questions were the same for both groups, except for additional items for EG to assess the ES’ bar charts and star rating elements.

We employed direct and indirect metrics to evaluate credibility. 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, results for believability, and reliability will 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 questions addressed the 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

Table 1 Results for the route selection scenario[3]

Characteristic	$\overline{CG}$	$\overline{EG}$	*p*
credibility	4.38	4.48	.490
believability*	4.1	4.57	.026
reliability	4.00	4.34	.160
mental_load	1.29	1.57	.162
info_amount*	2.76	3.29	.003
influence**	2.33	3.95	<.001

concerning the amount of cognitive load or information could be rated from *too less* (1) to *too much* (5).

8.1 Results

The study was conducted with 42 participants, equally distributed between groups. 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 applied as a co-variant in all variance tests to compensate for non-uniform distribution.

8.1.1 Route Selection

The results of variance analysis 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 also perceived significantly higher in EG (p=< .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 the ES' additional interface elements. The star rating was perceived helpful ($\overline{EG}$=4.09, σ=.94) and comprehensible ($\overline{EG}$=4.52, σ=.75) while creating only low cognitive load ($\overline{EG}$=1.48, σ=.75). The bar charts were considered comprehensible for time ($\overline{EG}$=4.05,σ=.89) and length comparison ($\overline{EG}$=4.15, σ=.88). They also positively impact believability of system recommendations, with time charts having a higher impact ($\overline{EG}$=4.15, σ=.75) than length charts ($\overline{EG}$=4.10, σ=.64). We conclude that facilitating comparability improves ANS credibility, because system recommendations are easier to validate.

8.1.2 Dynamic Traffic Messages

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 clear statement is possible. 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

[3] * = 5% significance level; ** = 0.1% significance level

Table 2 Results for the dynamic traffic information scenario[3]

Characteristic	$\overline{CG}$	$\overline{EG}$	*p*
credibility	3.43	4.19	.055
believability*	3.38	4.24	.017
reliability	3.48	4.14	.450
mental_load	2.71	2.71	.563
info_amount*	2.67	3.52	.017
influence	3.52	4.19	.127
choice_bel	3.95	3.71	.309

Table 3 Results for the deviation from route scenario[3]

Characteristic	$\overline{CG}$	$\overline{EG}$	*p*
credibility	4.09	4.57	.101
believability	3.71	4.29	.078
reliability*	3.62	4.33	.023
mental_load	2.81	2.19	.281
info_amount	2.95	3.05	.474
influence**	2.00	4.33	<.001
acoustic*	3.38	2.00	.002

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 leads to better informed decisions, which translate to less frustration and higher experienced credibility.

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

8.1.3 Deviation from Route

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 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.

8.2 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 the subjective nature of these characteristics. We presumed the subjectivity issue and therefore measured not just credibility but also related concepts which enable inferences for credibility.

In order to analyze general effects of these characteristics independent of specific scenarios, we formed combined scales in which we combined items from all scenarios that measure the same variable (e.g., credibility). Table 4 gives results of the variance analysis

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

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.

9 Conclusions

In this paper, we showed that ANS credibility can be improved by enhancing system usability. Consistent application of optimized interaction concepts enhances credibility and reduces *false/expected* messages. We proposed three major concepts for ANS interaction design: (1) providing choices when decisions are required, (2) providing relevant information which facilitates comparison of alternatives, and (3) enabling bidirectional communication to let drivers convey intention. As a result, system decisions are easier verifiable and drivers feels involved and supported in navigation-related decisions.

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]. The lab study provided unified conditions across groups, which simplified comparative evaluation of the developed concepts. As a drawback, actual risk of bad decisions experienced while driving is missing, which may impact credibility. Further studies with ANS that fully implement our concepts are required to analyze the effect of external factors on credibility. Long-term driver studies could also provide insights on how credibility and cognitive load develop over time.

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. The 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 wheel 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.

Acknowledgements This work was supported by Transregional Collaborative Research Centre SFB/TRR 62 (“Companion-Technology for Cognitive Technical Systems”) funded by the German Research Foundation (DFG). The authors would like to thank Roger Walk and Tina Seufert for their support and fruitful discussions.

References

1. Amditis, A., Andreone, L., Pagle, K., Markkula, G., Deregibus, E., Rue, M.R., Bellotti, F., Engelsberg, A., Brouwer, R., Peters, B., De Gloria, A.: Towards the automotive HMI of the future: overview of the AIDE-integrated project results. Trans. Intell. Transport. Sys. **11**, 567–578 (2010). DOI http://dx.doi.org/10.1109/TITS.2010.2048751. URL http://dx.doi.org/10.1109/TITS.2010.2048751
2. Bach, K.M., Jaeger, M.G., Skov, M.B., Thomassen, N.G.: Interacting with in-vehicle systems: understanding, measuring, and evaluating attention. In: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology, BCS-HCI ’09, pp. 453–462. British Computer Society, Swinton, UK, UK (2009). URL http://portal.acm.org/citation.cfm?id=1671011.1671070
3. Burnett, G.: The requirement for landmarks in vehicle navigation systems. The Journal of Navigation **53**(03), 499–510 (2000). DOI 10.1017/S0373463300001028
4. Chang, J.C., Lien, A., Lathrop, B., Hees, H.: 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, New York, NY, USA (2009). DOI http://doi.acm.org/10.1145/1620509.1620535. URL http://doi.acm.org/10.1145/1620509.1620535
5. Chittaro, L.: Visualizing information on mobile devices. Computer **39**(3), 40–45 (2006). DOI 10.1109/MC.2006.109
6. Fogg, B.J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., Treinen, M.: What makes web sites credible?: A report on a large quantitative study. In: CHI ’01: SIGCHI Conf. on Human factors in computing systems, pp. 61–68. ACM, New York, NY, USA (2001). DOI http://doi.acm.org/10.1145/365024.365037
7. Fogg, B.J., Soohoo, C., Danielson, D.R., Marable, L., Stanford, J., Tauber, E.R.: How do users evaluate the credibility of web sites?: A study with over 2,500 participants. In: DUX ’03: Conf. on Designing for user experiences, pp. 1–15. ACM, New York, NY, USA (2003). DOI http://doi.acm.org/10.1145/997078.997097
8. Fogg, B.J., Tseng, H.: The elements of computer credibility. In: CHI ’99: SIGCHI Conf. on Human factors in computing systems, pp. 80–87. ACM, New York, NY, USA (1999). DOI http://doi.acm.org/10.1145/302979.303001
9. Green, P., Levison, W., Paelke, G., Serafin, C.: Preliminary human factors design guidelines for driver information systems. Tech. Rep. UMTRI-93-21, University of Michigan Transportation Research Institute (UMTRI) (1993)
10. Hipp, M., Schaub, F., Kargl, F., Weber, M.: Interaction Weaknesses of Personal Navigation Devices. In: AutomotiveUI ’10: 2nd Intl. Conf. on Automotive User Interfaces and Interactive Vehicular Applications. ACM (2010)
11. Jensen, B.S., Skov, M.B., Thiruravichandran, N.: 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, New York, NY, USA (2010). DOI http://doi.acm.org/10.1145/1753326.1753517. URL http://doi.acm.org/10.1145/1753326.1753517
12. Kantowitz, B.H., Hanowski, R.J., C. Kantowitz, S.: Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. Human Factors **39** (1997)
13. Kaufmann, C., Risser, R., Geven, A., Sefelin, R.: Effects of simultaneous multi-modal warnings and traffic information on drivers behaviour. In: HUMANIST Proceedings 2008 (2008)
14. Khattak, A.J., Schofer, J.L., Koppelman, F.S.: Commuters’ enroute diversion and return decisions: Analysis and implications for advanced traveler information systems. Transportation Research Part A: Policy and Practice **27**(2), 101–111 (1993). DOI http://dx.doi.org/10.1016/0965-8564(93)90066-T
15. Lavie, T., Oron-Gilad, T., Meyer, J.: Aesthetics and usability of in-vehicle navigation displays. Int. J. Hum.-Comput. Stud. **69**, 80–99 (2011). DOI http://dx.doi.org/10.1016/j.ijhcs.2010.10.002. URL http://dx.doi.org/10.1016/j.ijhcs.2010.10.002
16. Lee, W.C., Ma, M.C., Cheng, B.W.: Field comparison of driving performance using a portable navigation system. The Journal of Navigation **63**(01), 39–50 (2010). DOI 10.1017/S0373463309990221
17. McKnight, D.H., Kacmar, C.J.: Factors and effects of information credibility. In: ICEC ’07: 9th Intl. Conf. on Electronic commerce, pp. 423–432. ACM, New York, NY, USA (2007). DOI http://doi.acm.org/10.1145/1282100.1282180

18. Muir, B.M.: Trust between humans and machines, and the design of decision aids. Intl. Journal of Man-Machine Studies **27**(5-6), 527–539 (1987). URL http://dx.doi.org/10.1016/S0020-7373(87)80013-5
19. Pauzi, A.: A method to assess the driver mental workload: The driving activity load index (DALI). IET Intelligent Transport Systems **2**(4), 315 –322 (2008). DOI 10.1049/iet-its:20080023
20. Rieh, S.Y., Danielson, D.R.: Credibility: A multidisciplinary framework. Annual review of information science and technology **41**, 307–364 (2007)
21. Robins, D., Holmes, J.: Aesthetics and credibility in web site design. Information Processing and Management **44**(1), 386–399 (2008). DOI http://dx.doi.org/10.1016/j.ipm.2007.02.003
22. Ross, T., Burnett, G.: Evaluating the human-machine interface to vehicle navigation systems as an example of ubiquitous computing. Intl. Journal of Human-Computer Studies **55**(4), 661–674 (2001). DOI http://dx.doi.org/10.1006/ijhc.2001.0495
23. Rydström, A., Grane, C., Bengtsson, P.: Driver behaviour during haptic and visual secondary tasks. In: AutomotiveUI '09: 1st Intl. Conf. on Automotive User Interfaces and Interactive Vehicular Applications, pp. 121–127. ACM, New York, NY, USA (2009). DOI http://doi.acm.org/10.1145/1620509.1620533
24. Tseng, S., Fogg, B.J.: Credibility and computing technology. Commun. ACM **42**(5), 39–44 (1999). DOI http://doi.acm.org/10.1145/301353.301402
25. Ware, C.: Information Visualization, 2. edn. Morgan Kaufmann (2004)