Unveiling the Lack of Scalability in Research on External Communication of Autonomous Vehicles

Mark Colley
Institute of Media Informatics
Ulm University, Germany
mark.colley@uni-ulm.de

Marcel Walch
Institute of Media Informatics
Ulm University, Germany
marcel.walch@uni-ulm.de

Enrico Rukzio
Institute of Media Informatics
Ulm University, Germany
enrico.rukzio@uni-ulm.de

Abstract
The traffic system is a complex network with numerous individuals (e.g., drivers, cyclists, and pedestrians) and vehicles involved. Road systems vary in various aspects such as the number of lanes, right of way, and configuration. With the emergence of autonomous vehicles, this system will change. Research has already addressed the missing communication possibilities when no human driver is needed. However, there is no common evaluation standard for the proposed external communication concept with respect to the complexity of the traffic system. We have therefore investigated the evaluation of these in Virtual Reality, in monitor-based, and in prototypical setups with special regard to scalability. We found that simulated traffic noise is a non-factor in current evaluations and that involving multiple people and multiple lanes with numerous vehicles is scarce.

Author Keywords
Self-driving vehicles; autonomous vehicles; scalability; external communication; interface design.

CCS Concepts
- General and reference → Surveys and overviews;
Introduction
Autonomous vehicles (AVs) are expected to change traffic enormously. Increased safety and facilitated mobility are some of the anticipated benefits of AVs [17]. AVs will be able to drive without any human driver present. Therefore, interpersonal communication such as hand gestures and eye-contact between drivers and pedestrians will be altered or vanish. Recent research projects aim to overcome this lack of communication through external communication modalities such as displays [19], LED strips [19, 29], movement patterns [51], projections [1] auditory or tactile cues [31] as well as combinations thereof [31] and enhancement of the infrastructure [44]. While scalability is mentioned as a potential problem in some of these publications [15, 28, 30, 31, 46], research has not yet addressed this issue specifically. Scalability in this context addresses the ability of the external communication concept to be used in scenarios with varying numbers of vehicles and/or pedestrians. We highlight this issue by conducting a thorough literature analysis with special regard to scalability factors (number of pedestrians, vehicles, lanes, noise). Therefore, we analysed publications with respect to their evaluations in the field of external communication of AVs. We found that some of these factors have been addressed (number of vehicles), that some of the imaginable scenarios are not relevant for investigation (one lane with one person and multiple vehicles) and that the following factors have not yet been properly addressed: noise, multiple pedestrians, and multiple lanes with vehicles.

Crossing Decisions
According to Rasouli and Tsotsos [42], there are 38 factors that influence crossing decisions. These can be broadly categorized into environmental and pedestrian factors. Focusing on scalability, especially the traffic characteristics of the environmental factors (Traffic volume, vehicle size, vehicle type, vehicle color) and the pedestrian factors are relevant: Group size, pedestrian flow, imitation as well as social status and social norms. Colley et al. [10] highlighted the need to include these factors in VR studies on external communication concepts.

Traffic Characteristics
Traffic characteristics include speed, flow, density, and time-space [33]. These vary greatly dependent on the location, e.g., in a city or on highways.

Traffic Volume
Various statistics are available documenting for example the traffic through the Alp tunnels in Italy (2014: on average 26,885 per day [47] or congestion in major cities (e.g., in Belfast, Great Britain [2] people spend approximately 200h per year in traffic jams). Traffic flow density, defined as vehicles per km (veh/km, sometimes additionally reported per lane), was reported for example in 1959 by Greenberg for Lincoln tunnel and Merrit Parkway [21]. It ranged from 20 to 165 veh/km. Treiber and Helbing modeled traffic with a varying density between 2 and 2400 veh/km [48].

Noise
Traffic noise is an important factor when evaluating auditory communication concepts. Fiedler et al. [18] report noise level in a Brazilian city regularly exceeds 65 dB(A). The Department of Transport Welsh Office published the calculation of road traffic noise [38], which includes multiple factors such as traffic flow, percentage heavy vehicles, traffic speed, gradient, road surface, and barriers. The New Zealand Transport Agency provides an online tool to calculate the noise level with the described
Parameters\(^1\). Parris and Schneider report values from 31 to 75 dB(A) for Mornington Peninsula [40].

**Analysis of External Communication Evaluation**

We evaluated the concepts reported by Colley et al. [9] and Rouchitsas and Alm [43] that provide a timely and thorough analysis over various conferences and scientific databases. Additionally, we screened the *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* for publications as a premier venue for automotive related publications. This resulted in a final sample of 38 publications. 13 studies were undergone in VR, 7 with a monitor (also online), and 9 are based on a hardware prototype.

**Exclusion Criteria**

Nine publications on external communication of AVs were excluded as they only propose concepts without evaluating them [15,19,35,44], report on design studies [46], do not give information on the road setup [14,50], or because they evaluated the concept in a non-realistic setting (e.g., in a paired comparison forced choice task in VR [26]). The publication by Dey et al. [16] was excluded because their focus laid on investigating gaze behavior while crossing to be able to design for external communication but without an actual external communication concept. Moore et al. [36] report on their experiment that investigated the use of an *implicit* external communication concept, therefore, this work was included in the analysis.

Some setups are crossed out, for details, see Section *Unexplored Setup.*


<table>
<thead>
<tr>
<th>lanes</th>
<th>vehicles</th>
<th>people</th>
<th>publications + [references]</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>1 [45]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>5 (one) [12,13,23,28,37]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>6 (one) [5,11] (two) [3,8,24]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>multiple</td>
<td>1 (multiple) [39]</td>
</tr>
</tbody>
</table>

*Table 1: Categorization based on VR study setup.*

<table>
<thead>
<tr>
<th>lanes</th>
<th>vehicles</th>
<th>people</th>
<th>publications + [references]</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>4 [1,4,6,20]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>3 (one) [27,41,49]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>0</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>multiple</td>
<td>0</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>multiple</td>
<td>0</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>multiple</td>
<td>0</td>
</tr>
</tbody>
</table>

*Table 2: Categorization on monitor study setup.*

<table>
<thead>
<tr>
<th>lanes</th>
<th>vehicles</th>
<th>people</th>
<th>publications + [references]</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>3 [29,31,51]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>2 (one) [25,32]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>one</td>
<td>1 (two) [7]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>multiple</td>
<td>1 [34]</td>
</tr>
<tr>
<td>one</td>
<td>multiple</td>
<td>multiple</td>
<td>0</td>
</tr>
<tr>
<td>multiple</td>
<td>multiple</td>
<td>multiple</td>
<td>2</td>
</tr>
</tbody>
</table>

*Table 3: Categorization on hardware study setup.*
**Evaluation Criteria**
The dimensions investigated were number of lanes, number of vehicles, and number of pedestrians. Each dimension was divided into the levels one and multiple. While we found that in some setups multiple lanes were designed (e.g., in VR), mostly only one lane was simulated. We refrained from adding another column to the tables, but distinguish this in the references (i.e., in a new line, (one) means that one lane was simulated in the following publications). In line with Rouchitsas and Alm, we divided the evaluation between a Virtual Reality (VR), a monitor based, and a hardware setup [43]. Additionally, we focused on reported noise levels, but do not report these in the tables as no reported noise was found.

**Findings from Literature Survey**
The overview given in Table 1, Table 2, and Table 3 shows that scalability is not properly addressed in neither of the three categories VR, monitor, or hardware.

Three setups have not been explored in either category: one lane with multiple vehicles and one person (highlighted lightgray), multiple lanes with one vehicle and multiple people (highlighted cyan), and one lane with multiple vehicles and multiple people (highlighted orange). While Merat et al. [34] evaluated the setup one lane with one vehicle and multiple people, this was not yet addressed in VR or in monitor-based studies (highlighted purple). This setup seems widely congruent to the setup one lane with multiple vehicles and multiple people, only differing in the number of vehicles. The multiple lanes with multiple vehicles and one person setup was mostly explored in VR (6/13; 46.15%) but is unexplored in monitor-based studies and was only addressed in one publication with an actual prototype [7].

Three publications (7.90%) report that multiple pedestrians were involved [22,36] or simulated [30].

**Discussion**
The overview has shown that several dimensions of traffic are not evaluated sufficiently with current concepts. It has also shown that auditory factors are mostly not taken care of.

**Unexplored Setup**
Several setups are unexplored. The setup one lane with multiple vehicles and one person seems unnecessary as it seems reasonable to assume that only the first vehicle halting will communicate with a person [8], therefore this variant is crossed out in all tables.

**Involving Multiple People**
Only 7.90% percent of the publications include multiple (simulated) people. According to Rasouli and Tsotsos, however, group size is a relevant factor for crossing decisions [42]. There could be several reasons to this: (1) For a first evaluation, simpler scenarios could be beneficial for participants to get used to. (2) Simulating participants brings a multitude of variation into a simulation. How should each person’s behavior and appearance be modeled? (3) In a real-world setting as reported in [22,36], it is not possible to control all variables as a public space was chosen.

**Noise**
We have not found explicit mentions of simulated noise. Mahadevan et al. [30] even mention that “ambient noise which we did not include in our current OnFoot scenarios, could also drown out auditory cues.” Therefore, they acknowledge these limitations, however, besides mentioning “street infrastructure” to provide auditory cues, do not go further into detail. Little focus on noise reporting could
also stem from little research into auditory external communication concepts (cf. [9]).

**Conclusion & Future Work**

38 publications in the field of external communication of AVs with regard to scalability aspects of the evaluation in VR, monitor-based, and real-world settings have been analyzed. We have found that, while scalability is an important factor, there is currently no best practice in evaluating this aspect. In general, there seems to be a lack of knowledge on how to best evaluate external communication concepts of AVs. However, there might not be a best practice as specific use cases for this communication have to be found and then be evaluated with the potential unique requirements in mind [36] We therefore strongly argue that evaluations should focus on more realistic evaluations, e.g., as done by Mahadevan et al. [30]. In the future, we plan to evaluate currently proposed concepts with special regard to scalability and provide best practices for the evaluation of external communication concepts with respect to scalability.

**Acknowledgements**

This work was conducted within the project 'Interaction between automated vehicles and vulnerable road users' (Intuitiver) funded by the Ministry of Science, Research and Arts of the State of Baden-Württemberg.

**REFERENCES**


[Online; accessed 23-DECEMBER-2019].


