

# Human Interpretation of Goal-Directed Autonomous Car Behavior

**Veronika Petrovych (veronika.petrovych@vti.se)**

Swedish National Road and Transport Research Institute (VTI) &  
Dept. of Computer & Information Science, Linköping University

**Sam Thellman (sam.thellman@liu.se)**

Dept. of Computer & Information Science  
Linköping University

**Tom Ziemke (tom.ziemke@liu.se)**

Dept. of Computer & Information Science, Linköping University &  
School of Informatics, University of Skövde

## Abstract

People increasingly interact with different types of autonomous robotic systems, ranging from humanoid social robots to driverless vehicles. But little is known about how people interpret the behavior of such systems, and in particular if and how they attribute cognitive capacities and mental states to them. In a study concerning people's interpretations of autonomous car behavior, building on our previous research on human-robot interaction, participants were presented with (1) images of cars – either with or without a driver – exhibiting various goal-directed traffic behaviors, and (2) brief verbal descriptions of that behavior. They were asked to rate the extent to which these behaviors were intentional and judge the plausibility of different types of causal explanations. The results indicate that people (a) view autonomous car behavior as goal-directed, (b) discriminate between intentional and unintentional autonomous car behaviors, and (c) view the causes of autonomous and human traffic behaviors similarly, in terms of both intentionality ascriptions and behavior explanations. However, there was considerably lower agreement in participant ratings of the driverless behaviors, which might indicate an increased difficulty in interpreting goal-directed behavior of autonomous systems.

**Keywords:** autonomous cars; self-driving; human-robot interaction; folk psychology; human-robot interaction; attribution; behavior explanation

## Introduction

The current state of technological development of autonomous cars suggests they are soon to become an everyday reality. Research on what can be expected from the introduction of autonomous vehicles to public roads is needed in order to ensure safety. One commonly voiced expectation is a decrease in road fatalities, both among car passengers and other road users. Vissers et al. (2017) recently provided detailed statistics showing that approximately 25% of road fatalities are so-called vulnerable road users (VRUs, i.e. pedestrians, cyclists, etc.). Most VRU fatalities happen in urban traffic and involve a collision with a passenger car. VRUs could – at least in theory – potentially benefit significantly from the introduction of autonomous vehicle technology and therefore must be taken into account in planning and

introducing changes in urban traffic environments. It is therefore important to note that in practice the introduction of autonomous cars in mixed traffic environments will not automatically lead to safer traffic. It might, for example, cause more misinterpretations among other road users – both VRUs and human drivers – when it comes to the interpretation of autonomous cars' behavior and the underlying cognitive mechanisms, e.g. the attribution of goals, beliefs and other mental states to such systems. At this point, however, relatively little is known about how people interpret the behavior of autonomous vehicles – or autonomous robotic systems in general (cf. Thellman et al., 2017a) – and in particular if and how they attribute cognitive capacities and mental states to such systems. The work presented here is a step in that direction.

## Related Work

Concerns have been voiced recently over how autonomous cars will be able to deal with traffic scenarios that traditionally require human communication (e.g. Nilsson et al., 2015). Brooks (2017), for example, discussed autonomous cars' inability to understand road user intentions communicated through body language. Soudi (2018) questioned how future autonomous vehicles will interpret flashing headlights, honks, and other forms of communication that people interpret differently depending on a variety of factors such as the time of day, the type of road, weather conditions, and intra-individual differences. These concerns focus on the technological side, whereas considerably less attention has been given to the people interacting with these systems. Road users generally rely on others to communicate their intentions, and such reliance on non-verbal human communication might cause difficulties for the observer in deciphering autonomous cars' intentions. Björklund et al. (2005) defined how drivers base their expectations of other road users on three aspects: traffic rules, design of the road, and the behavior of other road users. Moreover, past experiences in similar traffic situations influence how behavior is interpreted (Renner & Johansson, 2006). However, in response to the introduction of autonomous cars in mixed traffic environments people

may change their approach to handling traffic encounters. One reason for this is that people might not be able to easily discriminate between autonomous and manually driven cars.

The common assumption that autonomous cars will behave as human drivers does not actually hold for the less mature technologies being introduced nowadays. In urban traffic environments cars with higher levels of automation give higher priority to avoiding collisions rather than clear intention indication. That means, what the car will do next might be unclear for other road users even if it effectively maneuvers through the city traffic. A first attempt at investigating one-on-one interaction between an autonomous vehicle (emulated in this case, using a Wizard of Oz methodology) and a pedestrian was made by Rothenbuecher et al. (2016). This study intended to provide answers regarding how VRUs react towards a driverless car and how they interact with it without human contact. It is widely considered that road user interaction is defined by signals and behavior exhibited by both actors. While there may be a lot of reliance on active signals, such as eye-contact, passive signals as well as driving and walking behavior contribute highly to interaction and understanding between agents in dynamic traffic environments.

A recent study by Habibovic et al. (2016) looked at pedestrians and their understanding of automated vehicles' intentions as well as further actions when a driver is in the car, but does not appear to be driving. Rodriguez et al. (2017) compared how differences in autonomous car appearance influence pedestrians' decision to cross the street in an urban environment. The authors concluded that the autonomous cars were accepted by pedestrians and presented no perceived danger to observers. However, the absence of a driver inside the vehicle, compared to a present but visibly distracted driver, influenced trust levels and participants' decision to interact with the vehicle.

Difficulties detecting the intentions and goals that determine the behavior of autonomous cars may result in road accidents and other negative in-traffic experiences. Different approaches with external interfaces or road design used to communicate intentions have been tested, however many questions still remain open. It is therefore important to understand how people interpret the behavior of autonomous cars more generally. This might allow misinterpretations to be mitigated through, for example, the mindful design of autonomous vehicle appearance and behavior, altering the traffic environment accordingly, or by providing VRUs with relevant information about autonomous vehicles' limitations and capabilities.

## Method

### Methodology & Research Questions

In the research presented here, we considered the observer perspective on the behavior of autonomous versus manually driven cars, drawing on theory and methodology from literature on attribution and behavior explanation (for an overview, see Malle, 2004). We adopted the specific

methodology developed by Thellman et al. (2017a, 2017b)<sup>1</sup> in a human-robot interaction context, using a recent model of folk-psychological causal explanation of human behavior (Böhm & Pfister, 2015) as a starting point. Grounded in attribution theory, this model specifies seven cognitive categories assumed to be used by people in both behavior encoding and behavior explanation: goals, intentional actions, action outcomes, temporary states, dispositions, uncontrollable events, and stimulus attributes. Based on this framework, we compared participants' interpretations of car behaviors with and without a driver, to answer three specific questions: (Q1) How similar are people's judgments of the intentionality of driverless vs. human traffic behaviors? (Q2) How similar are people's judgments of the causes of driverless vs. human traffic behaviors? (Q3) How much do people agree in their judgments of driverless vs. human traffic behaviors?

### Participants & Design

Forty-nine individuals participated in the study ( $M_{(\text{age})} = 27$ ,  $SD = 5.7$  years;  $M_{(\text{self-assessed technical competence})} = 4.94$  out of 7,  $SD = 1.23$ ; 10% without driver's license, 15% possession for less than two years, 75% possession for more than two years; 22 women, 25 men, 2 unspecified gender). Participants were invited to complete a survey concerning traffic behavior. A majority were recruited at the Linköping University campus using a convenience sampling strategy. The only precondition for participation was self-assessed proficiency in the Swedish language. Each participant was assigned to one of two experimental conditions in a between-subjects study design: one group was given images of a car in various traffic scenarios with a human driver visibly placed behind the wheel. The other group was given images of a car without a human driver in more or less identical traffic scenarios (cf. Figure 1). In both conditions, the images were accompanied with brief written descriptions of the traffic behaviors. All 49 participants completed the questionnaire (26 in the human driver behavior condition, 23 in the autonomous condition). Distributions of age, gender, self-assessed technical competence and number of years with a driver's license did not significantly differ between experimental conditions. Out of the 23 participants in the autonomous condition who were presented with the additional question "Did you notice that the car was driverless/self-driving?", 17 selected the option "yes", 4 selected "no", 1 selected "do not know", and 1 did not answer.

---

<sup>1</sup> The present study is part of the second author's PhD thesis project with the general long-term motivation to understand how, when and why people take the *intentional stance* (Dennett, 1989) in interaction with different types of autonomous systems.

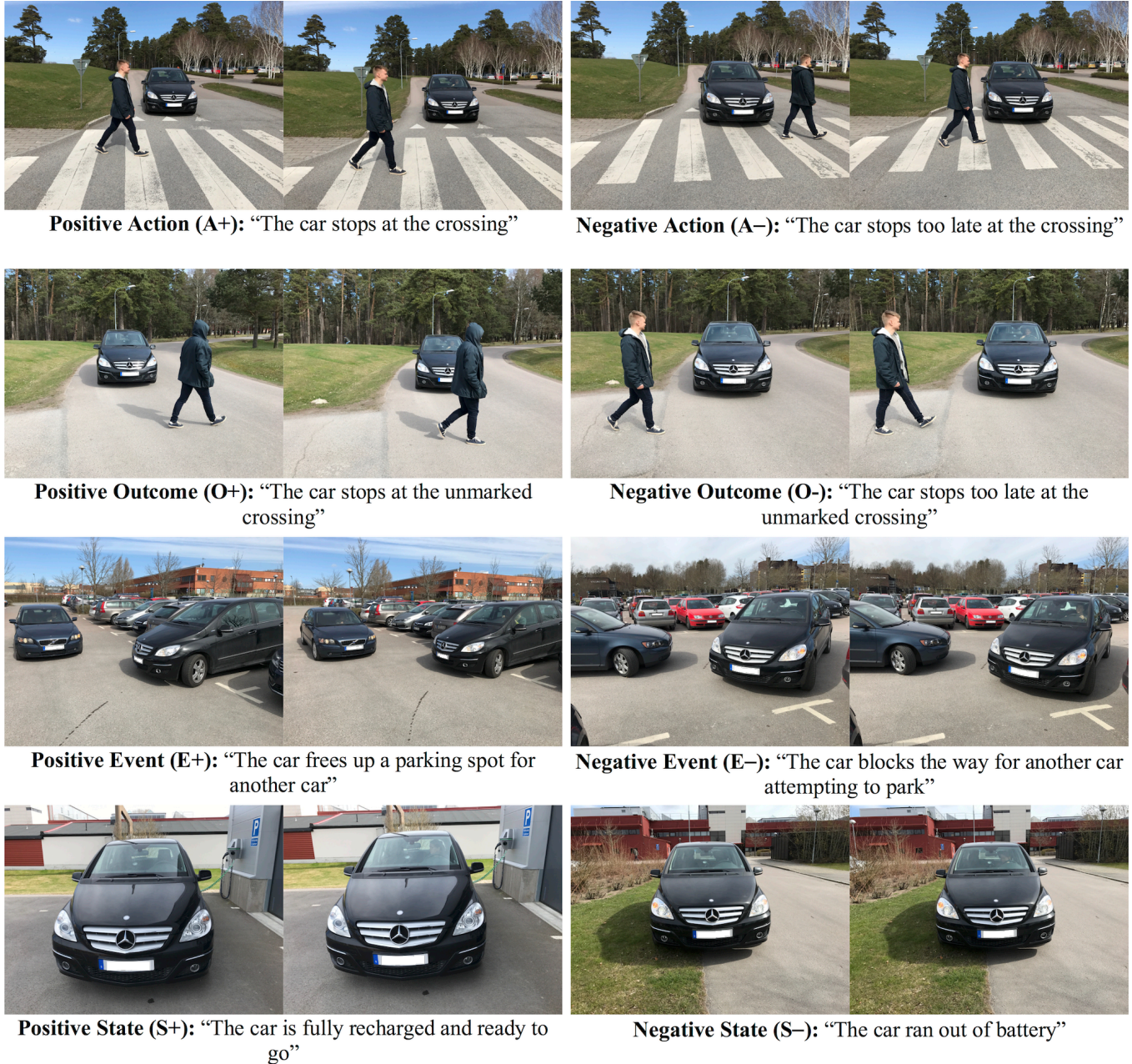


Figure 1: Descriptions (translated from Swedish) and images of driverless (left in each pair) versus human-driven (right in each pair) traffic behaviors used as experimental stimuli. Vehicle registration plates were blurred out post experiment.

### Stimuli

The aim of the present study was to explore people’s interpretations of traffic behavior exhibited by driverless versus human-driven cars. For the purpose of the experiment, we wanted to generate a relatively broad stimulus material covering a variety of traffic behaviors likely to occur in urban traffic environments. Accordingly, we based our stimuli selection on the four behavior types in Böhm and Pfister’s (2015) *Causal Explanation Network* (CEN) model: *actions*, *outcomes*, *events*, and *temporary states*. Böhm and Pfister concluded, from a series of experiments that involved participants sorting different

behaviors according to their similarity, that “the categories postulated in the CEN model cover the cognitive concepts that are relevant in the lay theory of behavior” (Böhm & Pfister, 2015, p. 8). We also introduced positively and negatively valenced variations of the four behavior types to account for positivity bias (cf. Thellman et al., 2017a). This resulted in the eight behavior types illustrated in Figure 1, with scenarios reflecting some of the challenges that human road users might face when encountering an autonomous car. Four of the eight behaviors illustrated in Figure 1 featured a scenario involving a car giving way to a VRU at a marked or unmarked crossing. Two behaviors involved a car interacting with another car at a parking lot. The remaining

two behaviors were related to the charging state of the car and did not involve interaction with other road users. It is important to note that we were not concerned with the representativeness of the selected behaviors as “good examples” of their corresponding behavior types; our purpose was primarily to select a reasonably broad and varied stimulus material with traffic situations likely to occur in urban environments.

### Instruments & Measures

The questionnaire was presented on a 9.7-inch (246 mm) screen sized tablet. Each behavior stimulus was presented on a separate page in a web browser window, with the image and description of the behavior displayed visibly on the top of the page and the questions below. The only difference between conditions was the exposure to images of an in-traffic vehicle with the presence (human condition) versus absence (driverless condition) of a human driver inside. The eight questionnaire pages were presented to participants in pseudo-randomized order to balance the influence of potential confounds related to prolonged task and stimuli exposure.

The questionnaire consisted of ten questions for each of the eight behavior stimuli illustrated in Figure 1 (resulting in 80 questions in total). The first three questions concerned the intentionality, controllability and desirability of the presented behaviors, and the following seven were concerned with judgments of the plausibility of various causes (including reasons) as explanations for the behaviors. The first series of questions were given in the form “Rate the extent to which the car’s behavior is X” where X was *intentional*, *controllable*, and *desirable*, respectively. The subsequent series of questions were given in the form “Rate how plausible it is that the cause of the car’s behavior is X”, where X was a *conscious goal*, an *action*, an *outcome*, an *uncontrollable event*, a *temporary state* (psychological or physical), a *disposition* and an *attribute* of someone or something in the car’s environment<sup>2</sup>. As mentioned above, the seven explanation types were derived from the CEN model (Böhm & Pfister, 2015). All ten questions were presented with the Likert-style option to select one of seven ordinal values ranging from “not at all” to “completely”.

## Results

### Q1: Intentionality

There was no statistically significant overall difference in participants’ ratings of the behaviors as *intentional* between human ( $M = 4.32, SD = 2.31$ ) and driverless ( $M = 4.42, SD$

<sup>2</sup> The Swedish terms used were: (for questions 1–3:) ”Uppskatta i vilken utsträckning bilens beteende är avsiktligt / kontrollerbart / önskvärt”; (for questions 4–10:) ”Uppskatta hur troligt det är att orsaken till bilens beteende är ett medvetet mål / en handling / resultatet av en handling / ett tillfälligt tillstånd (antingen psykologiskt eller kroppsligt) / en personlighetsegenskap / en egenskap hos någon eller någonting i bilens omgivning”.

$= 2.35$ ) conditions,  $t(382.704) = -.452, p = .651$ . Negative behaviors (–) were in general seen as more intentional when enacted by the driverless car ( $M = 3.71, SD = 2.27$ ) than the human driver ( $M = 2.97, SD = 1.86$ ),  $t(176.462) = -2.469, p = .014, d = 0.37$ . Ratings of positive behaviors (+) did not significantly differ between conditions,  $t(180.563) = 180.563, p = .080$ . See Figure 2 for an overview of ascribed level of intentionality to each of the eight behaviors in human and driverless conditions.

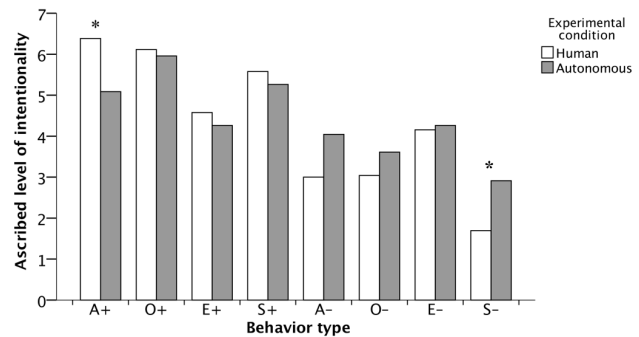


Figure 2: Interpretations of human and autonomous driving behaviors as intentional. See Figure 1 for spelled-out versions of the x-axis labels. \*Denotes a statistically significant difference ( $p < 0.05$ ).

### Q2: Causes

Participants were asked to rate the extent to which the behaviors enacted by the two types of agents were plausibly explained by seven different types of causes. The ratings were similar for human and driverless behaviors in 55 out of 58 individual cases, as assessed using multiple independent samples t-tests (Table 1). The effect sizes of the three differential effects were: *temporary state* as explanation for *positive state* (S+),  $d = 0.73$ ; *event* as explanation for *negative action* (A–),  $d = 0.63$ ; *goal* as explanation for *negative state* (–S),  $d = 1.18$ . The former two effect sizes might be classified as “moderate” and the latter effect as “large”.

### Q3: Agreement in ratings

Two-way random intra-class correlation coefficients (ICCs) with an absolute agreement definition were computed as a measure of the agreement in participant ratings within each experimental condition. Lower and higher 95% confidence interval bounds are reported in brackets. We report ICCs for participants’ *average* ratings (i.e., across all eight behaviors) of (firstly) behavior as intentional and (secondly) each of the seven behavior explanations as plausible.

The average level of agreement in participants’ interpretation of behavior as *intentional* was .96[0.90, 0.99],  $F(7, 175) = 27.83, p < .0005$ , in the human condition and .80[0.53, 0.93],  $F(7, 147) = 4.89, p < .0005$ , in the driverless condition (first pair of bars in Figure 3).

Table 1: Mean ratings of the plausibility of explanation types (rows) for human (left value in cell) and autonomous (right value in cell) driving behaviors (columns, see Figure 1 for spelled-out versions of the labels).

|                | A+        | O+        | E+        | S+               | A-               | O-        | E-        | S-               |
|----------------|-----------|-----------|-----------|------------------|------------------|-----------|-----------|------------------|
| Goal           | 5.8 / 4.7 | 5.8 / 5.8 | 4.9 / 4.4 | 5.4 / 5.9        | 3.9 / 4.2        | 3.0 / 3.6 | 4.2 / 4.5 | <b>1.4 / 2.9</b> |
| Action         | 6.2 / 5.4 | 6.3 / 6.3 | 6.1 / 5.6 | 5.2 / 5.9        | 4.9 / 5.3        | 5.3 / 5.7 | 6.0 / 5.7 | 2.8 / 4.0        |
| Outcome        | 6.5 / 6.2 | 6.4 / 6.1 | 6.0 / 5.9 | 6.2 / 6.4        | 5.9 / 6.1        | 5.9 / 5.7 | 6.0 / 6.3 | 5.5 / 5.9        |
| Event          | 2.5 / 2.3 | 2.9 / 2.1 | 3.4 / 2.7 | 2.1 / 2.7        | <b>3.1 / 2.2</b> | 3.7 / 3.1 | 2.5 / 2.7 | 2.9 / 2.7        |
| Temp. state    | 3.6 / 4.2 | 3.6 / 3.4 | 3.6 / 3.1 | <b>4.3 / 2.8</b> | 3.6 / 4.0        | 3.9 / 3.4 | 4.6 / 3.6 | 4.0 / 3.1        |
| Disposition    | 3.5 / 3.4 | 3.4 / 3.6 | 2.8 / 2.7 | 2.7 / 3.0        | 4.0 / 3.5        | 3.7 / 3.4 | 4.5 / 3.7 | 2.9 / 2.7        |
| Stimulus attr. | 4.6 / 4.6 | 5.3 / 4.7 | 3.8 / 4.4 | 4.2 / 4.3        | 4.1 / 4.3        | 4.9 / 4.7 | 4.4 / 4.9 | 3.2 / 3.9        |

Note. Statistically significant differences between groups ( $p < .05$ ) are marked in inverted colors.

The average level of agreement in judgments concerning the plausibility of the seven different behavior causes was as follows: *goal*, .94[0.87, 0.99],  $F(7,175) = 19.16, p < .0005$ , in the human condition and .83[0.59, 0.96],  $F(7, 154) = 5.73, p < .0005$ , in the driverless condition; *action*, .92[0.82, 0.98],  $F(7, 175) = 15.55, p < .0005$ , in the human condition and .68[0.29, 0.92],  $F(7, 154) = 3.52, p < .005$ , in the driverless condition; *outcome*, .50[-0.01, 0.87],  $F(7, 175) = 2.33, p < .05$ , in the human condition and .15[-0.66, 0.77],  $F(7, 154) = 1.22, p = .293$ , in the driverless condition; *event*, .63[0.21, 0.91],  $F(7, 175) = 3.16, p < .005$ , in the human condition and .13[-0.74, 0.77],  $F(7, 154) = 1.18, p = .319$ , in the driverless condition; *temporary state*, .20[-0.25, 0.73],  $F(7, 175) = 1.47, p = .180$ , in the human condition and .37[0.01, 0.79],  $F(7, 154) = 2.48, p < .05$ , in the driverless condition; *disposition*, .66[0.32, 0.91],  $F(7, 175) = 4.30, p < .0005$ , in the human condition and .23[-0.13, 0.72],  $F(7, 154) = 1.73, p = .107$ , in the driverless condition; *stimulus attribute*, .64[0.27, 0.91],  $F(7, 175) = 3.68, p < .005$ , in the human condition and -.10[-0.66, 0.62],  $F(7, 154) = 0.86, p = .543$ , in the driverless condition. The above ICC values are represented in Figure 3 (bar pairs 2–8). Values range from “poor” ( $< 0.5$ ), “moderate” (0.5–0.75), “good” (0.75–0.9), to “excellent” ( $> 0.9$ ) agreement.

In seven out of eight cases the average level of agreement in participants’ ratings was lower in the driverless condition than the human driver condition. Hence, agreement was systematically lower in the driverless condition.

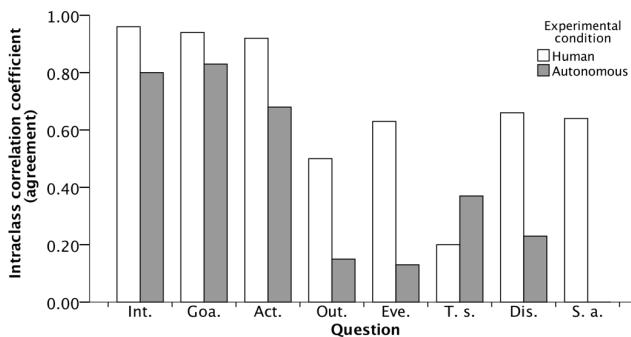


Figure 2: Average level of agreement in ratings of the intentionality (first bar pair) and plausibility judgments of causal explanations (second to eighth bar pair) of human and autonomous driving behaviors. See “Instruments & Measures” for spelled out versions of the x-axis labels.

## Discussion & Conclusions

The results show substantially similar ratings between human driver and driverless conditions on average, both in terms of ascribed intentionality and judged plausibility of behavior explanations. This is a clear indication that people view autonomous car behavior as goal-directed, discriminate between intentional and unintentional autonomous car behaviors, and view the causes of autonomous and human traffic behaviors similarly.

However, there was considerably lower agreement in participant ratings of the driverless car behaviors. This might be taken as an indication that – although understood similarly on average – some individuals may experience an increased difficulty in interpreting goal-directed autonomous car behavior. Agreement was particularly low in ratings concerning how the autonomous car behaviors were affected by persons or things in the environment of the vehicle, uncontrollable events, dispositions on part of the vehicle itself, and outcomes of the vehicle’s own behavior.

In order for the introduction of autonomous cars to have an overall positive impact on the traffic ecosystem it is important that all road users are able to understand autonomous vehicle traffic behavior to a degree comparable to their understanding of a human-driven vehicle behavior. This means that manufacturers will need to work actively on making sure that the behavior of autonomous cars is transparent and easy-to-understand for a broad variety of people with different needs and abilities (e.g., children on their way to school). The present study thus illustrates the need to (1) identify the group(s) of people that might experience increased difficulty interpreting autonomous cars (if any), and to (2) monitor how people’s interpretations and understanding change as road users become gradually familiarized with autonomous vehicles in traffic. These points are especially important to consider given that general traffic may very well soon be occupied by cars with a large variety of configurations and levels of automation – potentially making it even more difficult for road users to understand and interact with autonomous vehicles in traffic.

One limitation of the current study concerns the translation and meaning of the terms used in the questionnaire. It is possible that results would have been different if the questionnaire were given in another language

than Swedish. Furthermore, survey questions are always to some extent open to interpretation within subjects and different interpretation across subjects. While we do not take these limitations to be significant to the interpretation of our results, we think they warrant some caution in drawing conclusions based on ratings of individual questionnaire items.

When designing the stimuli for our experiment we have made an assumption of a fully autonomous vehicle being built upon an existing manufacturer's model with automation being an option rather than a must. Nevertheless, given the current trends and promises from car manufacturers, autonomous vehicles may or may not have passengers at all times of their operation and must obey the rules of traffic in any case.

Increasing vehicle automation in the city, where interaction with pedestrians and cyclists is crucial, has a potential to reduce the safety imbalance between road users that currently exists. People expect autonomous vehicles to follow written and unspoken rules, and to be able to recognize the intentions and actions of other road users in order to provide a safe transportation experience. The study presented here focuses on only a small fraction of what the future of autonomous vehicles in urban traffic holds. However, this is a very important aspect of road user interaction since interpretations of driving behavior play a crucial – but so far understudied – role in how future automated vehicles should be designed to communicate their intentions and thus integrate into the overall transportation system and traffic environment.

Last, but not least, we would like to emphasize that the contribution of this work lies not only in the concrete empirical results presented here. After all, autonomous vehicles, and human ways of interacting with them, will keep evolving continually over the next couple of decades, so these results might very well be very different a few years from now. At least equally important, in our opinion, is that our research (cf. also Thellman et al., 2017a) and related works (e.g. Habibovic et al., 2016) contribute to the incremental development of appropriate methodologies for the study of how humans interact with the increasing number and variety of autonomous technologies that are entering our society – and in particular how people attribute cognitive capacities and mental states to such systems. Cognitive science, we believe, has crucial contributions to make to the development and study of such interactive autonomous technologies.

### Acknowledgments

ST and TZ are supported by ELLIIT, the *Excellence Center at Linköping-Lund in Information Technology*. TZ is also supported by the Knowledge Foundation, Stockholm, under SIDUS grant agreement no. 20140220 (AIR, “*Action and intention recognition in human interaction with autonomous systems*”). The authors would also like to thank Ove Jansson, the person in the images, for his assistance.

### References

- Renner, L., & Johansson, B. (2006). Driver coordination in complex traffic environments. In *Proceedings of the 13th European conference on Cognitive ergonomics: trust and control in complex socio-technical systems* (pp. 35-40). ACM.
- Björklund, G.M. & Åberg, L. (2005). Driver behavior in intersections. Formal and informal traffic rules. *Transportation Research Part F*, 8(3), 239–253.
- Böhm, G., & Pfister, H. R. (2015). How people explain their own and others' behavior: A theory of lay causal explanations. *Frontiers in Psychology*, 6, 139
- Brooks, R. (2017). Robotic cars won't understand us, and we won't cut them much slack. *IEEE Spectrum*, 54(8), 34–51.
- Dennett, D. C. (1989). *The intentional stance*. MIT press.
- Habibovic, A., Andersson, M., Nilsson, M., Malmsten Lundgren, V. & Nilsson, J. (2016). Evaluating interactions with non-existing automated vehicles: Three Wizard of Oz approaches. *Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV) Gothenburg, Sweden*, 32–37.
- Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. MIT Press.
- Nilsson, M., Thill, S. & Ziemke, T. (2015). Action and intention recognition in human interaction with autonomous vehicles. *Proceedings of the CHI2015 workshop "Experiencing Autonomous Vehicles: Crossing the Boundaries between a Drive and a Ride"*, Seoul, Korea.
- Rodriguez Palmeiro, A., van der Kint, S., Vissers, L., Farah, H., de Winter, J., Hagenzieker, M.P. (2017) *Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment*. Unpublished manuscript.
- Rothenbacher, D., Li, J., Sirkin, D., Mok, B., Ju, W. (2016) Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. *25<sup>th</sup> IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 795–802.
- Soudi, A. (2018). Driverless cars might follow the rules of the road, but what about the language of driving? *The Conversation*. Retrieved from <http://www.theconversation.com>
- Thellman, S., Silvervarg, A., & Ziemke, T. (2017a). Folk-psychological interpretation of human vs. humanoid robot behavior: Exploring the intentional stance toward robots. *Frontiers in Psychology*, 8, 1962.
- Thellman, S., Silvervarg, A., & Ziemke, T. (2017b). Lay Causal Explanations of Human vs. Humanoid Behavior. In *International Conference on Intelligent Virtual Agents* (pp. 433-436). Springer, Cham.
- Vissers, L., Kint, S., Schagen, I., Hagenzieker, M. (2017). *Safe interaction between cyclists, pedestrians and automated vehicles. What do we know and what do we need to know?* R-2016-16, SWOV, The Hague.