SPAT: Situational Prosocial and Aggressive Behavior Perception in Traffic Scale

HATICE ŞAHİN IPPOLITI, Medieninformatik und Multimedia-Systeme, Universität Oldenburg, Germany
MARK COLLEY, UCL Interaction Centre, United Kingdom and Cornell Tech, U.S.
DEBARGHA DEY, Eindhoven University of Technology, The Netherlands
PHILIPP WINTERSBERGER, Interdisciplinary Transformation University, Austria
SHADAN SADEGHIAN, University of Siegen, Germany
ANDREAS LÖCKEN, Spiegel Institut, Germany
ANDRII MATVIIENKO, KTH Royal Institute of Technology, Sweden
AZRA HABIBOVIC, TRATON R&D, Sweden
HEIKO MÜLLER, Medieninformatik und Multimedia-Systeme, Universität Oldenburg, Germany
ANDREA HILDEBRANDT, Universität Oldenburg, Germany
SUSANNE BOLL, Medieninformatik und Multimedia-Systeme, Universität Oldenburg, Germany

Automated vehicles (AVs) reached technological maturity and will soon arrive on streets as traffic participants. Human traffic participants such as drivers, pedestrians, or cyclists will be increasingly confronted with the presence of AVs within their environment, not necessarily knowing or understanding what to expect and how to interact with them. Although AVs are designed to act safely, effective interaction in mixed traffic scenarios will depend on successful communication, interaction, or even negotiation beyond static rules and regulations. Prosocial behavior, such as yielding one's right of way, will be needed to resolve unclear traffic situations or foster traffic flow. However, what are the characteristics of such prosocial behavior, and how to measure this not only for automated vehicles but for all road users? Here, we describe a new scale to measure perceived social behavior in urban traffic scenarios. Through an online survey on $N = 318$ individuals and a validation study, we developed the Situational Prosocial and Aggressive Behavior in Traffic Scale and assessed it psychometrically.

CCS Concepts: • **General and reference**; • **Human-centered computing → HCI theory, concepts and models**; **Empirical studies in HCI**;

Additional Key Words and Phrases: Automated vehicles, social behavior in traffic, measurement methods for automated traffic, prosocial behavior, aggressive behavior, social perception of AVs, social perception in traffic, measurement of social behavior

Authors' Contact Information: Hatice Şahİn Ippoliti, hatice.sahin@uni-oldenburg.de, Medieninformatik und Multimedia-Systeme, Universität Oldenburg, Germany; Mark Colley, m.colley@ucl.ac.uk, UCL Interaction Centre, United Kingdom and Cornell Tech, U.S.; Debargha Dey, d.dey@tue.nl, Eindhoven University of Technology, Eindhoven, The Netherlands; Philipp Wintersberger, philipp.wintersberger@fh-hagenberg.at, Interdisciplinary Transformation University, Austria; Shadan Sadeghian, shadan.sadeghian@uni-siegen.de, University of Siegen, Germany; Andreas Löcken, andreas.loecken@thi.de, Spiegel Institut, Germany; Andrii Matviienko, andriim@kth.se, KTH Royal Institute of Technology, Sweden; Azra Habibovic, azra.habibovic@se.traton.com, TRATON R&D, Sweden; Heiko Müller, heiko.mueller@uni-oldenburg.de, Medieninformatik und Multimedia-Systeme, Universität Oldenburg, Germany; Andrea Hildebrandt, andrea.hildebrandt@uni-oldenburg.de, Universität Oldenburg, Germany; Susanne Boll, Medieninformatik und Multimedia-Systeme, Universität Oldenburg, Germany.

ACM Reference Format:
Hatice Şahİn Ippoliti, Mark Colley, Debargha Dey, Philipp Wintersberger, Shadan Sadeghian, Andreas Löcken, Andrii Matviienko, Azra Habibovic, Heiko Müller, Andrea Hildebrandt, and Susanne Boll. 2025. SPAT: Situational Prosocial and Aggressive Behavior Perception in Traffic Scale. In *17th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '25), September 21–25, 2025, Brisbane, QLD, Australia.* ACM, New York, NY, USA, 23 pages. https://doi.org/10.1145/3744333.3747812

1 Introduction

Humans are inherently social beings who interact daily in various contexts, such as work, commuting, and shopping, to communicate intentions, emotions, and needs. Social interaction extends beyond individual needs [60] and influences broader societal behavior. A key aspect of positive social interaction is "prosocial behavior," defined as *"[a] broad range of actions intended to benefit one or more people other than oneself – behaviors such as helping, comforting, sharing, cooperation, philanthropy, and community service"* [5, p. 243]. Traffic is one domain where social interactions frequently occur. While traffic rules provide structured guidance, unwritten norms shape many social interactions, such as yielding the right of way or using gestures of appreciation. In traffic, prosocial behavior is defined as *"driving behaviors that potentially protect the well-being of passengers, other drivers, and pedestrians, and that promote effective cooperation with others in the driving environment"* [21]. This includes both physical safety and psychological comfort, influencing factors like traffic flow and emissions. The emergence of automated vehicles (AVs) will alter social interactions in traffic, introducing mixed traffic environments with varying levels of automation [52]. While research has explored AV interactions with vulnerable road users (VRUs) [12, 23], fewer studies, such as Sadeghian et al. [51], have examined AV prosociality. Current traffic behavior metrics prioritize objective safety measures (e.g., time to collision), whereas social behaviors like courtesy and compromise remain challenging to quantify.

The goal of this study is to develop and validate a situational scale to measure social behavior perceptions of both human and automated road users in diverse traffic scenarios. We introduce SPAT (Situational Prosocial and Aggressive Behavior in Traffic Scale), a standardized questionnaire quantifying prosocial behavior in traffic interactions. We then test SPAT's usability in an AV-manual driver interaction scenario, demonstrating its external validity beyond human-to-human interactions. SPAT enables researchers to classify prosociality and aggression across diverse traffic contexts, aiding scenario development for AVs and analyzing driver behavior impacts. While several prosocial behavior scales exist, the unique contribution of our scale lies in its situational assessment approach, which captures context-specific variations in prosocial decision-making. Hence, in this study, we address a gap in the literature by developing a scale to measure social perception in traffic with a focus on situational specificity. Additionally, our research extends beyond human road users to include the social perception of automated vehicles, such as driverless cars and delivery robots, within mixed-traffic environments. Thus, our second objective is to create a versatile scale that evaluates the social behavior of both human and automated road users. The research questions explored in this work are: 1- How can we develop a reliable and valid scale to measure social perception in various traffic situations? 2- How can we develop an effective scale that measures the social behavior perception of both humans and automated vehicles?

2 Related Work

This section provides background on relevant research regarding the essentials of prosocial behavior in general and in the traffic context, followed by former approaches to measuring social behavior in traffic with self-assessment tools.

Prosocial Behavior in Traffic. Prosocial behavior involves interpersonal dynamics and various forms of helping acts [56]. Situational context and the recipient's perception shape prosocial behavior [13]. The empathy-altruism hypothesis [4] suggests that empathy triggers prosocial actions, with higher perceived need increasing helping likelihood. Alternatively, the arousal: cost–reward model [41] proposes that distress prompts prosocial acts, with individuals weighing the costs and benefits of helping. Lower costs of helping and higher perceived consequences of inaction increase prosocial behavior.

Prosocial behavior in traffic prioritizes the well-being and convenience of others. Examples include yielding to VRUs, allowing vehicles to merge, or pedestrians minimizing drivers' wait times at crossings. Previous research has explored prosocial and aggressive traffic behaviors [27]. Harris et al. [21] defined prosocial driving as actions that enhance safety and cooperation among road users. Eckoldt et al. [14] identified yielding to pedestrians as a common prosocial act, while Sikkenk and Terken [61] found that personal driving styles and recipient vulnerability influence right-of-way decisions. Efforts to promote prosocial driving have included storytelling and norm-activation [58]. Knobel et al. [29] introduced a character to encourage prosocial driving, while Wang et al. [68] proposed using like/dislike icons for driver feedback. Sadeghian et al. [51] investigated AV prosociality, showing that communication cues influence perceptions. Lanzer et al. [30] found polite AV communication increased trust across cultures. Colley et al. [9] showed bidirectional AV-pedestrian communication, such as thanking gestures, improved perceived intelligence.

Recent studies further highlight the role of prosocial behavior in traffic. Scott-Sharoni et al. [59] found post-ride feedback improved e-scooter rider interactions. Kim et al. [28] showed prosocial/asocial experiences affect micromobility users' well-being and behavior. Haué and Merlhiot [22] revealed that addressing uncivil AV behaviors enhances traffic flow. Escher et al. [16] examined Shared Automated Vehicles, suggesting passenger presence may trigger psychological effects like the bystander effect, influencing prosocial behavior.

Measuring Social Behavior in Traffic. Existing self-evaluation tools measure social behavior in traffic, primarily focusing on drivers. The Driver Behavior Questionnaire (DBQ) [43] assesses aberrant driving behaviors but lacks coverage of attitudes and emotions [64]. The Multidimensional Driving Style Inventory (MDSI) [66] extends this scope but remains centered on errors. Other scales, such as the Driver's Stress Profile (DSP) [31], Driving Behavior Inventory (DBI) [17, 19], and Aggressive Driving Behavior Scale (ADBS) [24], predominantly capture negative driving traits. To address this gap, Özkan and Lajunen [38] introduced the Positive Driver Behavior Scale (PDBS), which includes prosocial acts like politeness and environmental care. Some tools evaluate both prosocial and aggressive behaviors. The Prosocial and Aggressive Driving Inventory (PADI) [21] measures these dimensions separately, while Ward et al. [69] developed a similar scale focused on driver behavior toward cyclists. Pedestrian behavior has also been studied, often adapting driver-focused models [67]. The Pedestrian Behavior Scale (PBS) [18] and its derivative, the Pedestrian Behavior Questionnaire (PBQ) [11], have been widely used across various countries. Beyond self-assessment tools, researchers have explored social behavior perception in traffic using alternative frameworks. Sadeghian et al. [51] employed the Traffic Climate questionnaire [55], while Schwarting et al. [57] used the Social Value Orientation (SVO) Questionnaire [36] to analyze AVs' social behavior. However, most existing tools focus on general behaviors rather than situational specificity, limiting their application in dynamic traffic interactions [40]. As these scales do not assess how individuals perceive others' social behavior in specific scenarios—including interactions with AVs—we developed SPAT, exceeding these limitations. While initially designed to quantify AV perception, SPAT offers broad applicability across diverse road user interactions (e.g., cyclist-delivery robot, driver-e-scooter rider, or pedestrian-cyclist) and can facilitate self-evaluation in various situational contexts.

3 Study A: Creation of Situational Prosocial and Aggressive Behavior in Traffic Scale

The development and evaluation of SPAT followed a structured multistep process, incorporating expert input, pre-studies, and statistical validation We adhered to the guidelines by Boateng et al. [6] for scale creation and evaluation. We began with *item development* through workshops and linguist meetings to define the domain, generate items, and ensure content validity. Next, during *scale development*, we pre-tested questions for clarity and validated ground truth scenarios with 11 participants. A large online survey followed, involving item reduction and factor extraction with statistical expert input. For *scale evaluation*, we conducted confirmatory factor analysis to assess factorial validity. Finally, we tested external validity across different experimental conditions to ensure SPAT's applicability beyond the tested scenarios. Further details are provided in the following sections.

Pre-Survey Steps. A workshop by Sahin et al. [53] explored prosocial behavior in traffic, key characteristics, new interaction paradigms, and prosocial AV behavior. Eight AV researchers participated, highlighting the need for a quantifiable measure of situational prosocial behavior in traffic. Based on the workshop, SPAT was developed through a multistep approach. First, a group meeting defined prosociality, collecting related words rated for relevance (1–7 scale). Words scoring below 4 were excluded. To ensure versatility, semantic differentials were adopted (e.g., "mindless" vs. "mindful") rather than specific behaviors, enabling the scale to assess diverse road users and contexts, including AVs. The first author pre-clustered words into positive-negative adjective pairs, which were reviewed by a linguist with 13 years of expertise. After removing redundancies, 23 respondents rated the item pairs for relevance to prosocial and aggressive traffic behavior. Items scoring 4+ were refined further with expert input, reducing the list to 32 semantic differential items (e.g., "cautious — reckless"). Next, reference scenarios were defined and tested in a pre-study with 11 participants (M_{age} = 26.24, SD = 5.32; driving experience: M = 5.82 years, SD = 4.81). Participants evaluated road user behavior using the pilot survey and were interviewed on clarity. No issues were reported. They also rated scenarios for prosocial/aggressive perception and suggested refinements. The finalized ground truth scenarios included pedestrian-driver and driver-pedestrian interactions with both prosocial and aggressive behaviors (Figure 1, Appendix A). These scenarios ensured that item loadings reflected prosocial or aggressive behavior. After the pre-survey steps, we conducted an online vignette [1, 2] study with 318 participants from prolific.co that were pre-selected from the US to control the effect of cultural biases into a single culture. Finally, we validated SPAT in broader experimental conditions, including AV-human interactions, to test its sensitivity in diverse contexts (Section 4).

Vignette Survey Overview. An online within-subject vignette survey was conducted to evaluate SPAT, measuring perceived social behavior across all road users, including pedestrians, drivers, cyclists, and AVs. Ethical approval for both the online survey and the test track validation study was granted by the host institution's ethics committee. The study aimed to identify suitable items for prosocial and aggressive behavior and establish the scale's factorial structure. Given the absence of predefined facets, an exploratory factor analysis was conducted. Participants evaluated fictional situations using a within-subjects design, in other words, all participants rated all of the traffic scenarios.

Sample. A total of 318 participants from the USA (Age: M = 41.6, SD = 14.2, range: [19, 80]; Gender: 154 F, 153 M, 10 Non-binary, 1 Prefer not to tell) participated in the online survey. On average, they had a driving license for 11.10 years (SD = 15.65) and spent 2.48 hours daily on travel (SD = 6.07), including walking, cycling, driving, and public transport. Most drove their cars daily or multiple times a week. All participants had a high English proficiency (B2 = 4, C1 = 12, C2 = 302) and were recruited via Prolific, matching the study's demographic criteria.

Fig. 1. Scenarios categorized as prosocial and aggressive on the horizontal axis. Participant perspectives are divided on the vertical axis. Scenario 1: Prosocial Driver yielding the right of way to the participant from the pedestrians' perspective. Scenario 2: Aggressive driver taking the right of way of the participant at a pedestrian crossing. Scenario 3: A prosocial pedestrian yielding the right of way to the participant in the driver's perspective. Scenario 4: An aggressive pedestrian attempting to annoy the participant in the driver's perspective.

Procedure. The online survey was implemented via LimeSurvey and prolific.co. Participants could participate in the study through a computer, laptop, or tablet. Initially, they received a welcome message with details on the survey's duration and methodology. After providing consent, they completed demographic questions to confirm eligibility. Next, they were introduced to the general setting and the definition of prosocial behavior based on Harris et al. [21] (Appendix B).

Participants viewed four traffic scenarios in a pseudo-randomized order, each accompanied by descriptive images (see Figure 1). They responded to survey items presented in a pseudo-randomized order, evaluating the behavior of the other traffic participant from their perspective. Responses were recorded on a 7-point semantic differential scale (see Appendix B in Appendix for SPAT presentation and final set of items). Eight attention checks were included, and participants could provide open feedback for each situation and at the end. The survey, conducted in English, lasted about 15 minutes, and participants were compensated at a rate of $7.25 per hour.

Data Analysis. Content analysis and statistical methods were used to carefully select items. An exploratory factor analysis (EFA) was conducted on 50% of the sample, which was randomly split, to identify the number of dimensions. A confirmatory factor analysis (CFA) cross-validated the factorial structure using the second half of the sample. Additionally, a test track experiment in a bottleneck traffic scenario evaluated the external validity of the scale (see Section 4). Data analysis was performed with R 4.2.3 using up-to-date packages: `EFATools` version 0.4.4 [62], `psych` version 2.3.3 [48], and `eRm` version 1.0-2 [34]. Oblimin rotation was applied for factor interpretation [7], CFA was conducted with the `lavaan` package version 0.6-16 [49], and results were visualized with `semPlot` version 1.1.6 [15].

3.1 Study A Results

The results of the exploratory and confirmatory factor analyses are discussed in separate subsections below.

3.1.1 *Exploratory Factor Analysis.*

First, all inverted items were reverted by multiplying the answers by -1 and adding 8. This reverse-coding step ensured consistent contribution to the overall score; for example, inverted items such as unexpected - expected were adjusted to align with the regular format (expected - unexpected). Each scenario evaluation was then randomly split into two samples: exploration and confirmation. The analyses confirmed the data's suitability for factor analysis (Kaiser–Meyer–Olkin (KMO = .89) and Bartlett's test of sphericity ($\chi^2(496)$ = 3597.50, p<.001) [65]. Scree tests and parallel analysis using weighted least squares revealed the factor structure for each scenario: Scenario 1 (prosocial driver) had two factors, Scenario 2 (aggressive driver) had four, Scenario 3 (prosocial pedestrian) had three, and Scenario 4 (aggressive pedestrian) had three.

Upon inspection, the four-factor scenario had only one item loading onto one factor, leading to the conclusion that a three-factor structure was most consistent. Factor analysis was performed with oblimin rotation, as correlations among dimensions were expected [32], and polychoric correlations were used for ordinal data. A factor loading threshold of 0.45 was applied [10, 20]. The three-factor model demonstrated high internal consistency across scenarios (Total Sn1 ω = 0.95, Sn2 ω = 0.78, Sn3 ω = 0.94, Sn4 ω = 0.92). Recurring items across scenarios were prioritized for robustness, leading to the finalized factor structure (see Table 1). Without predefined expectations, factors were identified and named Socialness (S), Awareness (A), and Predictability (P) based on item content.

Table 1. Final Set of Items and Their Factor Loading Values (Pattern Matrices) on Each Scenario

	Sn1			Sn2			Sn3			Sn4		
Item	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
S1 Cooperative - Competitive	**0.972**			**0.598**	0.123	-0.541	**0.596**	0.376	-0.164	**0.740**		-0.216
S2 Helpful - Unhelpful	**0.564**	0.432		**0.948**	-0.125		**0.871**			**0.909**		
S3 Courteous - Impolite	**0.775**	0.150		**0.933**	-0.171		**0.616**	0.261		**0.942**	-0.178	
S4 Prosocial - Antisocial	**0.880**			**0.705**		-0.197	**0.600**	0.296	-0.121	**0.890**		-0.203
A5 Trustworthy Untrustworthy	0.349	**0.560**		0.750	0.160			**0.759**	0.358	0.709	0.194	0.106
A6 Mindless - Mindful (i)	0.605	0.282	0.184	0.226	**0.603**	0.351	0.694	0.122		0.234	**0.536**	0.321
A7 Ungrateful - Thankful (i)	0.219	0.410	0.119	0.385	**0.567**	0.159		**0.484**	0.551	0.753	0.122	
A8 Aware - Unaware	0.449	0.440		-0.375	**0.670**		0.895			-0.212	**0.831**	
A9 Inattentive - Attentive (i)	0.562	0.299			**0.802**		0.781	0.130	-0.141		**0.810**	
A10 Acknowledging - Ignoring	0.899		-0.235	-0.108	**0.841**	-0.124	0.553	0.259	-0.138	0.219	**0.700**	
P11 Unexpected - Expected (i)			**1.006**	0.329	0.141	0.246		-0.159	**0.716**			**0.536**
P12 Predictable - Unpredictable	-0.172	0.388	**0.562**	0.227	0.203	0.228	0.196		**0.590**	0.158	0.173	**0.574**

*(i) inverted item, retained items are marked in **bold**. Sn1: Prosocial driver, Sn2: Antisocial driver, Sn3: Prosocial pedestrian, Sn4: Antisocial pedestrian. Factor 1: Socialness, Factor 2: Awareness, Factor 3: Predictability*

3.1.2 *Confirmatory Factor Analysis.*

Twelve items met the exploratory factor analysis criteria (Table 1) and were examined in the confirmation sample (Table 3). Following Mair [33] (Section 4.3.2.), the Partial Credit Model (PCM) was applied to assess item fit. PCM was run per factor in each scenario, with items removed if they had low p-values in at least three scenarios and infit/outfit t-values beyond ±2. No items were excluded, resulting in a 12-item SPAT scale comprising four *Socialness*, six *Awareness*, and two *Predictability* items (Figure 5 in the appendix). A confirmatory factor analysis validated the three-factor model (Figure 2, Figure 6, Figure 7, Figure 8). Predictability item loadings were constrained for stability. For clarity, only Figure 2 (prosocial driver) is included in the main text, while others are in the appendix. Path diagrams display standardized factor loadings, latent variable correlations, and error variances.

Factor loadings in the prosocial driver scenario (Scenario 1) were strong ($0.63 \leq \lambda \leq 0.93$), with high correlation between Socialness and Awareness (λ = 0.96) but weaker correlations with Predictability and Socialness (λ = 0.23) and Predictability and Awareness (λ = 0.43). The aggressive driver scenario (Scenario 2) also showed strong item-factor relationships, except for A8 (λ = 0.45). Socialness and Awareness correlated moderately (λ = 0.63), while Awareness and

Predictability (λ = 0.48), and Socialness and Predictability (λ = 0.36) were weaker. In the prosocial pedestrian scenario (Scenario 3), factor loadings were strong ($0.67 \leq \lambda \leq 0.92$), with an almost perfect correlation between Socialness and Awareness (λ = 0.99). Weaker correlations between Awareness and Predictability (λ = 0.22), as well as Socialness and Predictability (λ = 0.10) were observed. The aggressive pedestrian scenario (Scenario 4) mirrored Scenario 2, with A8 showing a weaker loading (λ = 0.45) and moderate Socialness-Awareness correlation (λ = 0.73). Weaker correlations were further observed between latent factors Awareness and Predictability (λ = 0.51), and Socialness and Predictability (λ = 0.53).

Model fit was assessed using χ^2, CFI ($\geq$ 0.95), RMSEA (upper $\leq$ 0.10, lower $\leq$ 0.05), and SRMR ($\leq$ 0.08) (Table 4), per Mair [33] (Section 2.4.1., p. 44). Prosocial scenarios (Sn. 1 and 3) demonstrated strong fit (CFI $\geq$ 0.997, RMSEA upper $\leq$ 0.083, lower $\leq$ 0.031, SRMR $\leq$ 0.073). Aggressive scenarios (Sn. 2 and 4) showed poorer fit (CFI $\leq$ 0.935, RMSEA lower $\geq$ 0.136, upper $\geq$ 0.175, SRMR $\geq$ 0.136). Modification indices indicated that adding residual covariances (A8-A9, A9-A10) improved fit (Table 4). McDonald's Omega confirmed internal consistency (Table 2). Socialness showed the highest reliability in Scenario 1 and remained strong across scenarios. Awareness had consistently high values, particularly in Scenarios 1 and 3. Predictability varied, peaking in Scenario 2 but maintaining generally good reliability.

Table 2. Summary of Mcdonald's Omega (ω) Values by Scenario and Factors for Internal Consistency

Scenarios	**Socialness**	**Awareness**	**Predictability**
Sn1	0.869	0.875	0.799
Sn2	0.708	0.689	0.773
Sn3	0.839	0.878	0.697
Sn4	0.783	0.739	0.757

Lastly, to examine factor correlations across scenarios, we analyzed the full sample and estimated a model for each factor across all scenarios (Figure 3, Figure 9, Figure 10). For readability, Figure 3 (Socialness path diagram) is included in the main text, while the other two appear in the appendix. Correlations between Socialness across scenarios were weak ($-0.17 \leq \lambda \leq 0.15$). Similarly, Awareness correlations were weak ($-0.10 \leq \lambda \leq 0.04$), as were those for Predictability ($-0.10 \leq \lambda \leq 0.20$). These results indicate the situation specificity of the scale, effectively capturing social perception variations across the surveyed traffic scenarios.

Table 3. Descriptive statistics of survey scenarios for overall and sub-dimensions scores

Scenarios	Composite Average	Socialness	Awareness	Predictability
Sn1	6.07 (1.58)	6.69 (0.87)	6.46 (1.06)	3.66 (1.75)
Sn2	2.03 (1.65)	1.26 (0.77)	2.54 (1.96)	2.05 (1.32)
Sn3	6.03 (1.55)	6.63 (0.79)	6.46 (0.96)	3.56 (1.71)
Sn4	2.52 (1.75)	1.59 (0.99)	3.07 (1.97)	2.74 (1.47)

Average scores and standard deviations in parentheses.
Sn1: Prosocial driver, Sn2: Antisocial driver, Sn3: Prosocial pedestrian, Sn4: Antisocial pedestrian.

4 Study B: External Validation of SPAT

While human-to-human interaction is a key traffic-related application of SPAT, the rise of AVs introduces a new interaction context. Highly automated vehicles must share the road with manual drivers and VRUs, requiring appropriate behavior and communication in ambiguous situations. To evaluate this, SPAT was tested in a controlled test track experiment to assess its external and ecological validity. Further details and additional findings beyond SPAT are available in Şahin İppoliti et al. [54].

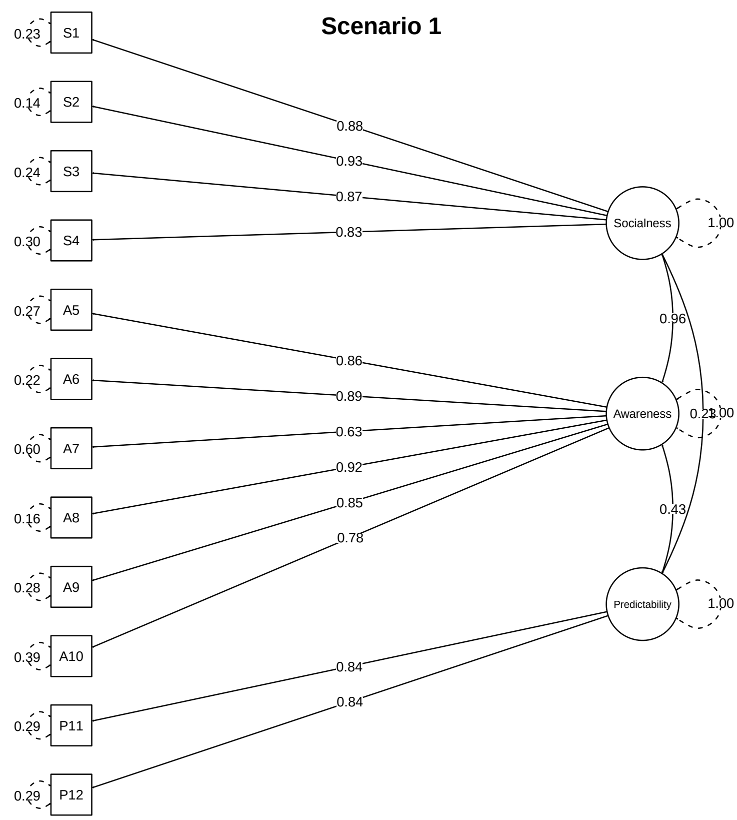


Fig. 2. Path diagram of Scenario 1. Participant from the pedestrian's perspective encounters a prosocial driver. Items are presented as squares with their associated error variances and latent variables as circles with their associated variances and covariance (edge). The values on the arrows are the loadings.

Table 4. Confirmatory Factor Analysis Fit Measures on Each Scenario

Scenarios	χ^2	df	p	CFI	RMSEA 90% CI lower	RMSEA 90% CI upper	SRMR
Sn1	79.03	51	0.007	0.997	0.031	0.083	0.073
Sn2	245.185 (167.784)	51 (50)	0 (0)	0.91 (0.945)	0.136 (0.102)	0.175 (0.143)	0.149 (0.115)
Sn3	62.981	51	0.121	0.999	0	0.067	0.064
Sn4	266.982 (114.07)	51 (50)	0 (0)	0.935 (0.981)	0.145 (0.068)	0.183 (0.112)	0.136 (0.085)

CFI: Comparative fit index, SRMR: Standardized root-mean-square residual, Sn1: Prosocial driver Sn2: Antisocial driver, Sn3: Prosocial pedestrian, Sn4: Antisocial pedestrian. Values in parentheses show the results of the model specifying additional covariances between the residuals of A8 and A9, and A9 and A10.

Study Overview. Twenty-four participants (8 female, 16 male, age [20–67], M = 30.21, SD = 13.44 years) drove their own cars through a naturally narrowed path with parked cars on both sides (Figure 4B). As they approached, a Wizard of Oz AV, controlled by a hidden driver [50], simultaneously reached the opposite side (Figure 4D). An LED matrix on the AV's radiator grille (Figure 4A) remained off (baseline) or signaled deceleration/acceleration with moving bars (Figure 4C). Since intuitiveness was not tested, participants were introduced to the interface beforehand. The ghost driver matched the AV's behavior to the displayed signal, maintaining an average speed of 12 km/h in the baseline

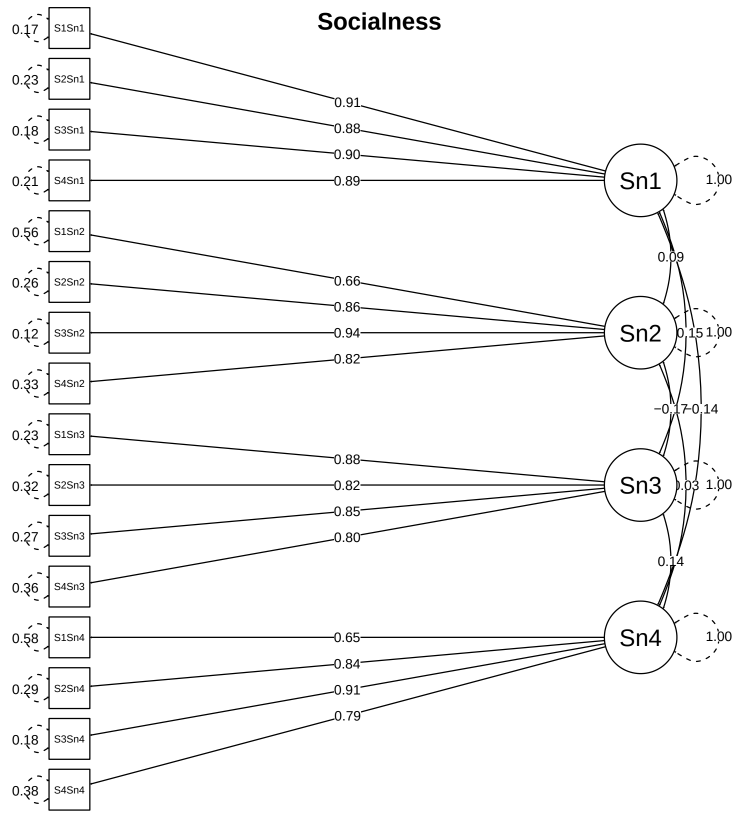


Fig. 3. Path diagram of factor Socialness and correlations of each scenario. Items are presented as squares with their associated error variances and latent variables as circles with their associated variances and covariance (edge). The values on the arrows are the loadings. Notation example: S1Sn1 indicates item S1 from Scenario 1.

unless adjusting to the participant's speed. They also ensured sufficient time and space for participants to decide who would cross first. After each of 9 pseudo-randomized encounters (3 repetitions per condition), participants rated the AV's perceived social characteristics using SPAT's exploratory 32-item set on a tablet.

Data Analysis. SPAT evaluations of the AV across three conditions were analyzed using linear mixed-effects models (LMM) with the `lmer` function in R package lme4 version 1.1-27.1 [3]. Composite SPAT scores and sub-dimensions were used as outcome variables, where higher scores indicated perceiving the AV as more prosocial (with inverted items for better interpretability). The results below are based on the 12 final items (Appendix B in Appendix). External interface conditions (eHMI Acceleration, eHMI Deceleration, Baseline) were modeled as fixed effects, while within-subject variance, sex, and age-related variability were included as random effect factors.

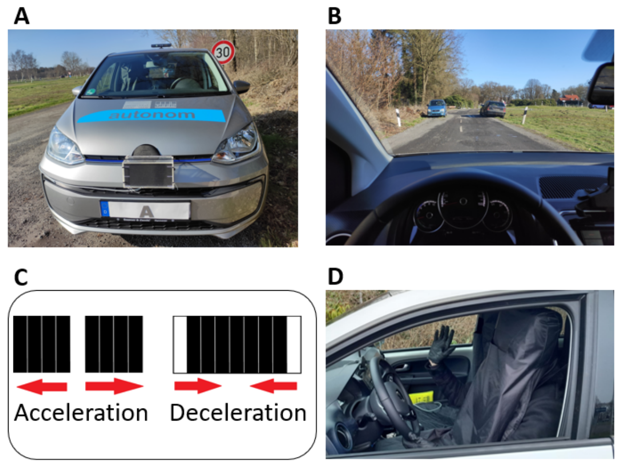


Fig. 4. **A**: Wizard of Oz AV used in the study. The external Human-Machine Interface (eHMI) in the form of an LED matrix is attached above the covered license plate. **B**: Participant perspective while approaching the bottleneck scenario. **C**: Acceleration and deceleration intention cues. The white bar extends from the middle to the sides on acceleration, while it merges in the middle on deceleration intention. **D**: Ghost driver hidden under a car seat costume.

Table 5. Driving choices and descriptive statistics of experimental conditions for each SPAT score

Conditions	# of times passed/waited	SPAT Composite Average	Socialness	Awareness	Predictability
Baseline	32/40	4.86 (0.92)	4.84 (1.09)	4.93 (0.88)	4.69 (1.01)
eHMI Acceleration	19/53	5.14 (0.68)	5.03 (0.84)	5.13 (0.67)	5.40 (0.85)
eHMI Deceleration	64/8	5.37 (0.74)	5.53 (0.79)	5.29 (0.71)	5.28 (1.07)

Average scores and standard deviations in parentheses.

4.1 Study B Results

Indicating deceleration intention significantly increased the likelihood of the AV being perceived as prosocial compared to the baseline condition ($\beta = 5.87$, $t(199.641) = 3.22$, $Pr(>|t|) < .01$). However, indicating acceleration intention did not lead to a meaningful change in prosocial perception compared to baseline ($\beta = 2.94$, $t(199.641) = 1.61$, $Pr(>|t|)$ $= .10$). Consistent with the composite scores, the Socialness dimension of SPAT showed significantly higher social perception scores when the AV indicated deceleration intention ($\beta = 2.77$, $t(199.607) = 3.95$, $Pr(>|t|) < .001$), while no significant difference was observed in the acceleration condition compared to baseline ($\beta = 0.76$, $t(199.607) = 1.08$, $Pr(>|t|) = .27$). Compared to the baseline (no eHMI), an AV indicating deceleration via eHMI significantly increased participants' perception of the AV's awareness ($\beta = 2.16$, $t(200.28) = 2.39$, $Pr(>|t|) < .05$). No significant difference was found for the acceleration intention condition ($\beta = 1.09$, $t(200.28) = 1.21$, $Pr(>|t|) = .22$). The Predictability dimension of SPAT indicated that participants found the AV's behavior significantly more predictable in both the deceleration eHMI condition ($\beta = 1.18$, $t(189.99) = 2.72$, $Pr(>|t|) < .01$) and the acceleration eHMI condition ($\beta = 1.41$, $t(189.99) =$ 3.26, $Pr(>|t|) < .01$) (see Table 6).

Table 6. Linear Mixed-Effects Model results of eHMI conditions on composite SPAT scores and SPAT sub-dimensions of Socialness, Awareness, and Predictability

Predictors	Composite Average	Socialness	Awareness	Predictability
(Intercept)	26.92 ***	13.45 ***	15.53 ***	8.27 ***
eHMI Acceleration	2.94	0.76	1.10	1.42 **
eHMI Deceleration	5.87 **	2.78 ***	2.17 *	1.18 **
Random Effects				
σ^2	119.43	17.76	29.52	6.78
τ_{00} ID	0.00	0.00	0.00	0.08
τ_{00} age	11.74	2.21	2.28	0.01
τ_{00} sex	0.00	0.00	0.00	0.33
N ID	24	24	24	24
N sex	2	2	2	2
N age	17	17	17	17
Observations	216	216	216	216
Marginal r^2 / Conditional r^2	0.046 / NA	0.072 / NA	0.026 / NA	0.051 / 0.106

*** $p < .001$, ** $p < .01$, * $p < .05$

Note: *LMM Estimates and Random Effects are reported for composite average scores and Socialness, Awareness, and Predictability dimension scores of SPAT.*

5 Discussion

This section discusses how SPAT addresses the need for an emerging evaluation tool, explores the impact of its sub-dimensions on individuals' judgments of prosocial and aggressive behavior, and provides further insights into construct and external validity. Additionally, the limitations of the current study and directions for future research are considered.

5.1 Situational Prosocial and Aggressive and Behavior in Traffic Scale (SPAT)

Existing social behavior in traffic scales are predominantly tailored for drivers and their general road user characteristics (e.g., PADI [21], DBQ [43], MDSI [66], DSP [31], DBI [19], ADBS [24]). Some scales have been developed for pedestrians (e.g., PBS [18], PBQ [11], SBP [35]) and cyclists [69]. While these tools are specific and useful when applied to their intended purpose, they have three key drawbacks that limit their ability to evaluate social behavior across different traffic situations.

Firstly, these evaluation tools focus on self-assessing the general social behavior of traffic participants. When used in experiments with different social conditions, they may not reveal situational differences across traffic conditions. For example, if a person answers a self-assessment questionnaire about their general driving characteristics in different traffic situations (e.g., highway driving or slow-moving traffic), they are likely to provide similar responses, as these tools are designed to assess general driving tendencies. However, in specific situations (e.g., when in a hurry or encountering road rage), individuals may deviate from their usual behavior. This limitation underscores the need for a situational assessment of social behavior in traffic, as highlighted by researchers in this area [53]. SPAT was designed to fill this gap by evaluating situational differences in how road users perceive social behavior in traffic, enabling an assessment of individuals' social receptivity to changes in behavior or vehicle design across different traffic scenarios. Secondly, many existing scales, including the ones mentioned above, are self-assessment tools not designed to understand how individuals perceive the behavior of other road users within a social context. SPAT, on the other hand, was intentionally developed with semantic differential items (e.g., reckless – cautious), avoiding first-person statements such as "I find cyclists reckless when they cut in front of me." This design allows SPAT to assess how individuals perceive the prosocial or aggressive behavior of other road users, rather than focusing solely on self-assessment. Thirdly, with the rise of micromobility, it is becoming increasingly important to expand the perspectives of traffic participants beyond just drivers, pedestrians, and cyclists. An inclusive assessment tool that can be used from any road user's perspective would significantly enhance the current body of research on social behavior in traffic. By formulating the items as adjectives rather than using specific statements (e.g., "Cyclists who cut in front of others are reckless"), SPAT remains versatile and applicable to a wide range of road users. This flexibility even allows SPAT to assess how individuals perceive the social behavior of non-human traffic participants, such as autonomous vehicles (AVs). The Situational Prosocial and Aggressive Behavior in Traffic Scale (SPAT) we developed serves as a valuable tool for studies focused on prosocial behaviors in traffic. By providing a standardized method for assessing such behaviors, it enables comparisons across various research contexts and contributes to a common framework for understanding how different factors influence prosocial and aggressive interactions in diverse traffic environments.

5.2 Key Ingredients for Prosocial Perception in Traffic: Socialness, Predictability, and Awareness

Upon closer inspection of the sources of perceptual differences regarding the social behavior of other road users, SPAT revealed that the general social outlook of road users, the predictability of their behavior, and their awareness of the impact of their actions on the situation help individuals form judgments about the prosocial and aggressive behavior of others in traffic.

Socialness. Social qualities such as cooperation and helpfulness are key in traffic judgments, even in this formal, regulated environment. Traffic, while governed by rules, is a social domain [8], where cues such as eye contact and gestures emerge, especially in ambiguous situations [63]. Existing scales, such as PDBS [38], already incorporate socialness by highlighting politeness and helpfulness in positive driving behavior. In any human environment, including traffic, prosocial behavior is crucial for fostering cooperation. Therefore, cooperation, helpfulness, and courtesy shape individuals' prosocial perceptions of others. This factor was consistently effective in describing road users' behavior across various scenarios, making it a strong measure of social perception.

Predictability. While traffic rules are essential, they don't fully capture prosocial behavior in traffic. Scales such as PADI [21] and DARQ [26] include rule-related items to assess self-evaluation of social driving. Predictable behavior is a key factor in judging whether a road user's actions are prosocial or aggressive. For example, feedback on the first scenario highlighted that while yielding the right of way is prosocial, following the rules is expected and resolves conflicts in a usual manner. Similarly, in the third scenario, a pedestrian giving way was seen as prosocial, but insisting on the right of way, though lawful, wouldn't be considered aggressive, as it's within traffic rules.

Awareness. SPAT analysis shows that Awareness items have more impact in aggressive scenarios than in prosocial ones. In prosocial behavior scenarios, Awareness and Socialness items are highly correlated, but this correlation doesn't harm model fit. However, item A9 (Attentive - Inattentive) loads onto Socialness in prosocial scenarios and onto Awareness in aggressive ones, which affected the fit in the CFA of aggressive scenarios. By adding covariances between A8, A9, and A10, we improved the model fit (Table 4). Awareness helps define aggressive behavior, but doesn't seem significant for prosocial behavior beyond general socialness. Consistent with Dovidio et al. [13], participants judged behavior within its social context. Sometimes the participant's intentions or the behavior's outcome played a bigger role in determining whether a road user was perceived as prosocial or aggressive. For example, participants noted that a prosocial driver's sudden stop could be hazardous, while an aggressive pedestrian, even when following rules, was judged harshly due to their perceived intent. This suggests that both behavioral intentions and outcomes influence social perception, with their importance varying by the situation.

5.3 Construct and External Validation of SPAT

We used four traffic scenarios with two perspectives and road user types to identify the best working items. Confirmatory factor analysis showed good fit for prosocial scenarios. For aggressive scenarios, adding covariances between A8, A9, and A10 improved fit, suggesting that removing item A9 (Inattentive - Attentive) could enhance the model, particularly for aggressive scenarios. SPAT showed sensitivity to situation specificity, maintaining discriminative power across different scenarios. We tested SPAT's external validity in a controlled test track study with a game-of-chicken [42] or bottleneck [44–47] scenario. In bottleneck situations, when the AV indicated deceleration via an external display, it was perceived as more social, aware, and predictable. Acceleration intention signaling also made the AV more predictable. In both cases, participants linked the AV's intention cues to *predictability* and perceived it as *aware* during deceleration. Deceleration, giving the human driver the right of way, was viewed as more prosocial, in line with Sadeghian et al. [51]. These results confirm that SPAT effectively detects prosocial perceptions of road users and is sensitive to situational differences. Given that the tool is developed in the context of AV technology, it is important to acknowledge that the underlying policies of AI agents can shape prosociality; situating our work within the broader HCI discourse on designing prosocial AI enhances the contextual validity and potential generalizability of our scale [25, 39].

5.4 Limitations and Future Directions

While the set of scenarios in our study may appear orchestrated, they serve as a controlled foundation for initial validation; importantly, the situational format of our scale holds strong potential for real-world applications, complementing recent advances in physiology-based prosociality assessment [37] by providing a structured yet adaptable tool for context-sensitive evaluation. In the future, SPAT can be further validated by applying it to various traffic situations and road user perspectives (e.g., AVs, e-scooter riders, delivery robots, pedestrians). An additional elicitation round may be needed for future work. This study presents the first scale for assessing prosocial behavior in traffic, laying the foundation for standardized traffic scales. While the validation was conducted outdoors with a real car, the ecological validity may still be low. Future research should explore its use in more realistic traffic scenarios, involving participants both inside and outside vehicles. SPAT was tested with a U.S. sample to avoid cultural bias, but testing in different populations may yield different results. Thus, this scale is primarily suited for English speakers in the U.S. and requires validation for other languages and regions. Furthermore, the cultural generalizability of the scale needs to be further evaluated, since prosocial behavior may change according to the cultural differences. Broader cultural testing and more realistic environments are needed for wider application.

6 Conclusion

We introduced SPAT, a new scale for measuring road users' prosocial and aggressive behavior in urban traffic scenarios. SPAT features semantically distinct items made of adjectives that capture various factors influencing the perception of prosocial and non-prosocial behaviors. The scale's factorial structure includes three dimensions: socialness, predictability, and awareness. Confirmatory factor analysis confirmed its construct validity, and external validity was demonstrated in an experimental study between drivers and a Wizard of Oz automated vehicle using a bottleneck scenario. SPAT offers a useful tool for researchers in traffic psychology, social interaction in traffic, and the study of social interactions with automated vehicles.

Acknowledgments

The authors thank all study participants and linguist expert Audrey Debije-Popson for their contributions.

References

[1] Herman Aguinis and Kyle J Bradley. 2014. Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational research methods* 17, 4 (2014), 351–371.

[2] Christiane Atzmüller and Peter M Steiner. 2010. Experimental vignette studies in survey research. *Methodology* 6 (2010), 128–138. doi:10.1027/1614-2241/a000014

[3] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1 (2015), 1–48. doi:10.18637/jss.v067.i01

[4] C. Daniel Batson. 2010. *Altruism in Humans*. Oxford University Press, USA. doi:10.1093/acprof:oso/9780195341065.001.0001

[5] C Daniel Batson. 2012. A history of prosocial behavior research.

[6] Godfred O. Boateng, Torsten B. Neilands, Edward A. Frongillo, Hugo R. Melgar-Quiñonez, and Sera L. Young. 2018. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health* 6 (2018), 15 pages. doi:10.3389/fpubh.2018.00149

[7] Jürgen Bortz and Christof Schuster. 2011. *Statistik für Human-und Sozialwissenschaftler: Limitierte Sonderausgabe*. Springer-Verlag, Cham.

[8] Barry Brown, Mathias Broth, and Erik Vinkhuyzen. 2023. The Halting Problem: Video Analysis of Self-Driving Cars in Traffic. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 12, 14 pages. doi:10.1145/3544548.3581045

[9] Mark Colley, Jan Henry Belz, and Enrico Rukzio. 2021. Investigating the Effects of Feedback Communication of Autonomous Vehicles. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21)*. Association for Computing Machinery, New York, NY, USA, 263–273. doi:10.1145/3409118.3475133

[10] Andrew L Comrey and Howard B Lee. 1992. A first course in factor analysis, 2nd edn. hillsdale, nj: L.

[11] Shuchisnigdha Deb, Lesley Strawderman, Janice DuBien, Brian Smith, Daniel W Carruth, and Teena M Garrison. 2017. Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the US population. *Accident Analysis & Prevention* 106 (2017), 191–201.

[12] Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. 2020. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives* 7 (2020), 100174.

[13] John F Dovidio, Jane Allyn Piliavin, David A Schroeder, and Louis A Penner. 2017. *The social psychology of prosocial behavior*. Psychology Press, East Sussex, United Kingdom.

[14] Kai Eckoldt, Marc Hassenzahl, Matthias Laschke, Thies Schneider, Josef Schumann, and Stefan Könsgen. 2016. The Gentleman. A prosocial assistance system to promote considerate driving. In *Proceedings on the 10th Conference on Design and Emotion*. Design & Emotion Society, The Netherlands, 307–314.

[15] Sacha Epskamp. 2015. semPlot: Unified visualizations of structural equation models. *Structural Equation Modeling: a multidisciplinary journal* 22, 3 (2015), 474–483.

[16] Bengt Escher, Jakob Peintner, and Andreas Riener. 2024. The Significance of the Bystander Effect on Personal Responsibility in Critical Situations in Shared Automated Vehicles. In *Proceedings of the 16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 384–394.

[17] AI Glendon, Lisa Dorn, Gerald Matthews, E Gulian, DR Davies, and LM Debney. 1993. Reliability of the driving behaviour inventory. *Ergonomics* 36, 6 (1993), 719–726.

[18] Marie-Axelle Granié, Marjorie Pannetier, and Ludivine Gueho. 2013. Developing a self-reporting method to measure pedestrian behaviors at all ages. *Accident Analysis & Prevention* 50 (2013), 830–839.

[19] E Gulian, G Matthews, Aleck Ian Glendon, DR Davies, and LM Debney. 1989. Dimensions of driver stress. *Ergonomics* 32, 6 (1989), 585–602.

[20] JF Hair jr, RE Anderson, RL Tatham, and WC Black. 1998. Multivariate Data Analysis (Fifth Eds. ed.).

[21] Paul B Harris, John M Houston, Jose A Vazquez, Janan A Smither, Amanda Harms, Jeffrey A Dahlke, and Daniel A Sachau. 2014. The Prosocial and Aggressive Driving Inventory (PADI): A self-report measure of safe and unsafe driving behaviors. *Accident Analysis & Prevention* 72 (2014), 1–8.

[22] Jean-Baptiste Haué and Gaëtan Merlhiot. 2024. How do Automated Vehicles Influence Other Road Users and Sometimes Elicit Uncivil Behaviors?. In *Proceedings of the 16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 374–383.

[23] Kai Holländer, Mark Colley, Enrico Rukzio, and Andreas Butz. 2021. A Taxonomy of Vulnerable Road Users for HCI Based On A Systematic Literature Review. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 158, 13 pages. doi:10.1145/3411764.3445480

[24] John M. Houston, Paul B. Harris, and Marcia Norman. 2003. The Aggressive Driving Behavior Scale: Developing a self-report measure of unsafe driving practices. *North American Journal of Psychology* 5, 2 (2003), 269–278. The purpose of this study was to investigate the reliability and validity of a new 11-item measure of aggressive driving, the Aggressive Driving Behavior Scale, which focuses on behaviors, rather than cognitions, emotions, or motivational states. Based on a sample of 211 undergraduates (111 women and 89 men), the study examined the convergent validity of

the new scale with measures of hostility, hypercompetitiveness, and aggressive thoughts and emotions experienced while driving. A principal component analysis of the Aggressive Driving Behavior Scale yielded two factors that form reliable subscales labeled Speeding and Conflict Behavior. As expected, the total scale and its two subscales correlated with hostility, hypercompetitiveness, as well as aggressive driving-related thoughts and emotions. The results suggest that the scale can be used as a research tool and a self-assessment instrument. (PsycINFO Database Record (c) 2016 APA, all rights reserved).

[25] Xinyue Hu, Kumar Akash, Shashank Mehrotra, Teruhisa Misu, and Mark Steyvers. 2024. Prosocial Acts Towards AI Shaped By Reciprocation And Awareness. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 46.

[26] Dongchao Huo, Jinfei Ma, and Ruosong Chang. 2023. The development and application of the drivers' attitudes of right-of-way questionnaire (DARQ). *Transportation research part F: traffic psychology and behaviour* 94 (2023), 67–82.

[27] Sherrie-Anne Kaye, David Rodwell, Natalie Watson-Brown, Chae Rose, and Lisa Buckley. 2022. Road users' engagement in prosocial and altruistic behaviors: A systematic review. *Journal of safety research* 82 (2022), 342–351.

[28] Sooyeon Kim, Shashank Mehrotra, Kumar Akash, Teruhisa Misu, and John D Lee. 2024. Prosociality Matters: How Does Prosocial Behavior in Interdependent Situations Influence the Well-being and Cognition of Road Users?. In *Proceedings of the 16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 395–404.

[29] Martin Knobel, Marc Hassenzahl, Simon Männlein, Melanie Lamara, Josef Schumann, Kai Eckoldt, Matthias Laschke, and Andreas Butz. 2013. Become a Member of the Last Gentlemen: Designing for Prosocial Driving. In *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces* (Newcastle upon Tyne, United Kingdom) *(DPPI '13)*. Association for Computing Machinery, New York, NY, USA, 60–66. doi:10.1145/2513506.2513513

[30] Mirjam Lanzer, Franziska Babel, Fei Yan, Bihan Zhang, Fang You, Jianmin Wang, and Martin Baumann. 2020. Designing communication strategies of autonomous vehicles with pedestrians: an intercultural study. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, USA, 122–131.

[31] John A Larson. 1996. *Steering clear of highway madness: A driver's guide to curbing stress and strain.* BookPartners, Sao Paulo, Brazil.

[32] Ruben D Ledesma, Pere J Ferrando, Mario A Trógolo, Fernando M Poó, Jeremías D Tosi, and Cándida Castro. 2021. Exploratory factor analysis in transportation research: Current practices and recommendations. *Transportation research part F: traffic psychology and behaviour* 78 (2021), 340–352.

[33] Patrick Mair. 2018. *Modern psychometrics with R.* Springer, Cham.

[34] P. Mair and R. Hatzinger. 2007. Extended Rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software* 20, 9 (2007), 1–20. https://www.jstatsoft.org/v20/i09

[35] Emilio Moyano Díaz. 1997. Teoría del Comportamiento Planificado e intención de infringir normas de tránsito en peatones. *Estudos de Psicologia (Natal)* 2 (1997), 335–348.

[36] Ryan O Murphy, Kurt A Ackermann, and Michel JJ Handgraaf. 2011. Measuring social value orientation. *Judgment and Decision making* 6, 8 (2011), 771–781.

[37] Abinay Reddy Naini, Zhaobo K Zheng, Teruhisa Misu, and Kumar Akash. 2025. Self-Supervised Learning-Based Multimodal Prediction on Prosocial Behavior Intentions. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.

[38] Türker Özkan and Timo Lajunen. 2005. A new addition to DBQ: Positive driver behaviours scale. *Transportation Research Part F: Traffic Psychology and Behaviour* 8, 4-5 (2005), 355–368.

[39] Ana Paiva, Filipa Correia, Raquel Oliveira, Fernando Santos, and Patrícia Arriaga. 2021. Empathy and prosociality in social agents. In *The handbook on socially interactive agents: 20 Years of research on embodied conversational agents, intelligent virtual agents, and social robotics volume 1: methods, behavior, cognition*. 385–432.

[40] Jean-Luc Patry. 2011. Methodological consequences of situation specificity: biases in assessments. *Frontiers in psychology* 2 (2011), 18.

[41] Jane Allyn Piliavin et al. 1981. *Emergency intervention.* Academic Press, Cambridge, Massachusetts, USA.

[42] Anatol Rapoport and Albert M Chammah. 1966. The game of chicken. *American Behavioral Scientist* 10, 3 (1966), 10–28.

[43] James Reason, Antony Manstead, Stephen Stradling, James Baxter, and Karen Campbell. 1990. Errors and violations on the roads: a real distinction? *Ergonomics* 33, 10-11 (1990), 1315–1332.

[44] Michael Rettenmaier, Deike Albers, and Klaus Bengler. 2020. After You?!–Use of External Human-Machine Interfaces in Road Bottleneck Scenarios. *Transportation research part F* 70 (2020), 175–190.

[45] Michael Rettenmaier and Klaus Bengler. 2021. The Matter of How and When: Comparing Explicit and Implicit Communication Strategies of Automated Vehicles in Bottleneck Scenarios. *IEEE Open Journal of Intelligent Transportation Systems* 2 (2021), 282–293. doi:10.1109/OJITS.2021.3107678

[46] Michael Rettenmaier, Sabrina Dinkel, and Klaus Bengler. 2021. Communication via motion–Suitability of automated vehicle movements to negotiate the right of way in road bottleneck scenarios. *Applied ergonomics* 95 (2021), 103438.

[47] Michael Rettenmaier, Moritz Pietsch, Jonas Schmidtler, and Klaus Bengler. 2019. Passing through the Bottleneck - The Potential of External Human-Machine Interfaces. In *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, Paris, France, 1687–1692. doi:10.1109/IVS.2019.8814082

[48] William Revelle. 2019. *psych: Procedures for Psychological, Psychometric, and Personality Research.* Northwestern University, Evanston, Illinois. https://CRAN.R-project.org/package=psych R package version 1.9.12.

[49] Yves Rosseel. 2014. The lavaan tutorial.

[50] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. 2015. Ghost driver: a platform for investigating interactions between pedestrians and driverless vehicles. In *Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular*

Applications. Association for Computing Machinery, New York, NY, USA, 44–49.

[51] Shadan Sadeghian, Marc Hassenzahl, and Kai Eckoldt. 2020. An Exploration of Prosocial Aspects of Communication Cues between Automated Vehicles and Pedestrians. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) *(AutomotiveUI '20)*. Association for Computing Machinery, New York, NY, USA, 205–211. doi:10.1145/3409120.3410657

[52] Taxonomy SAE. 2016. Definitions for terms related to driving automation systems for on-road motor vehicles. *SAE Standard J* 3016 (2016), 2016.

[53] Hatice Sahin, Heiko Müller, Shadan Sadeghian, Debargha Dey, Andreas Löcken, Andrii Matviienko, Mark Colley, Azra Habibovic, and Philipp Wintersberger. 2021. Workshop on Prosocial Behavior in Future Mixed Traffic. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. Association for Computing Machinery, New York, NY, USA, 167–170.

[54] Hatice Şahin İppoliti, Angelique Daudrich, Debargha Dey, Philipp Wintersberger, Shadan Sadeghian, and Susanne Boll. 2023. A Real Bottleneck Scenario with a Wizard of Oz Automated Vehicle-Role of eHMIs. In *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 280–290.

[55] Jens Schade, Lars Rößger, Johannes Eggs, Robert Follmer, and Bernhard Schlag. 2019. Entwicklung und Überprüfung eines Instruments zur kontinuierlichen Erfassung des Verkehrsklimas. Issue 289.

[56] David A. Schroeder and William G. Graziano. 2015. 3The Field of Prosocial Behavior: An Introduction and Overview. In *The Oxford Handbook of Prosocial Behavior*. Oxford University Press, Oxford, UK. arXiv:https://academic.oup.com/book/0/chapter/212096115/chapter-ag-pdf/44596024/book_28077_section_212096115.ag.pdf doi:10.1093/oxfordhb/9780195399813.013.32

[57] Wilko Schwarting, Alyssa Pierson, Javier Alonso-Mora, Sertac Karaman, and Daniela Rus. 2019. Social behavior for autonomous vehicles. *Proceedings of the National Academy of Sciences* 116, 50 (2019), 24972–24978.

[58] Shalom H Schwartz. 1977. Normative influences on altruism. In *Advances in experimental social psychology*. Vol. 10. Elsevier, Amsterdam, The Netherlands, 221–279.

[59] Sidney T Scott-Sharoni, Shashank Mehrotra, Kevin Salubre, Miao Song, Teruhisa Misu, and Kumar Akash. 2024. Can we enhance prosocial behavior? Using post-ride feedback to improve micromobility interactions. In *Proceedings of the 16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 294–304.

[60] Kennon M Sheldon, Andrew J Elliot, Youngmee Kim, and Tim Kasser. 2001. What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of personality and social psychology* 80, 2 (2001), 325.

[61] Marin Sikkenk and Jacques Terken. 2015. Rules of conduct for autonomous vehicles. In *Proceedings of the 7th international conference on automotive user interfaces and interactive vehicular applications*. ACM, New York, NY, USA, 19–22.

[62] Markus D. Steiner and Silvia Grieder. 2020. EFAtools: An R package with fast and flexible implementations of exploratory factor analysis tools. *Journal of Open Source Software* 5, 53 (2020), 2521. doi:10.21105/joss.02521

[63] Matus Sucha, Daniel Dostal, and Ralf Risser. 2017. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention* 102 (2017), 41–50.

[64] Nebi Sümer. 2003. Personality and behavioral predictors of traffic accidents: testing a contextual mediated model. *Accident Analysis & Prevention* 35, 6 (2003), 949–964.

[65] Barbara G Tabachnick, Linda S Fidell, and Jodie B Ullman. 2013. *Using multivariate statistics*. Vol. 6. Pearson, Boston, MA, USA.

[66] Orit Taubman-Ben-Ari, Mario Mikulincer, and Omri Gillath. 2004. The multidimensional driving style inventory—scale construct and validation. *Accident Analysis & Prevention* 36, 3 (2004), 323–332.

[67] Romane Vandroux, Marie-Axelle Granié, Mathilde Jay, Cédric Sueur, and Marie Pelé. 2022. The pedestrian behaviour scale: A systematic review of its validation around the world. *Accident Analysis & Prevention* 165 (2022), 106509.

[68] Chao Wang, Jacques Terken, Jun Hu, and Matthias Rauterberg. 2016. " Likes" and" Dislikes" on the Road: A Social Feedback System for Improving Driving Behavior. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, USA, 43–50.

[69] Nicholas J Ward, Kari Finley, Jay Otto, David Kack, Rebecca Gleason, and T Lonsdale. 2020. Traffic safety culture and prosocial driver behavior for safer vehicle-bicyclist interactions. *Journal of safety research* 75 (2020), 24–31.

A Scenario Descriptions

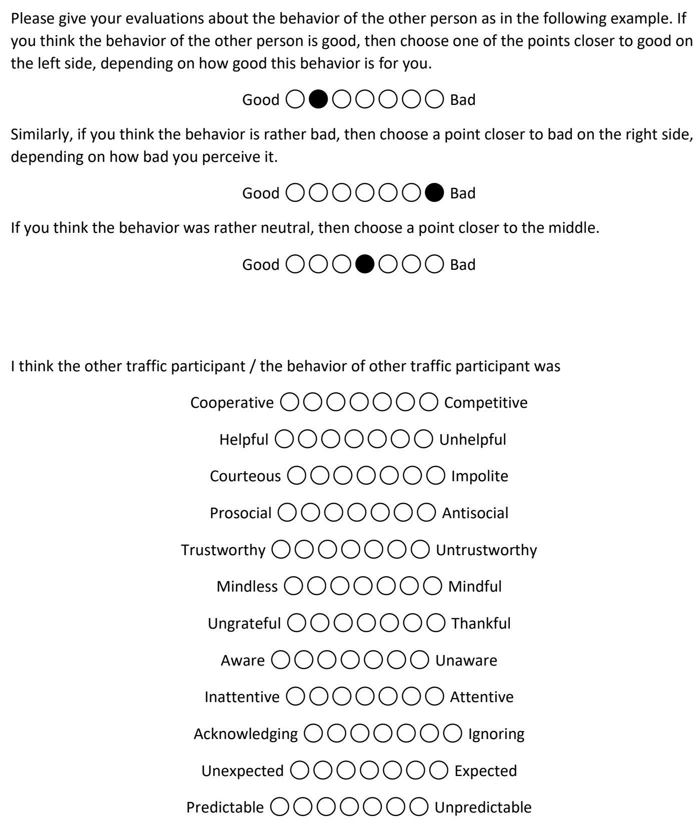
- Scenario 1: "Imagine you are a pedestrian, and you want to cross the street at a place where there aren't any traffic lights or a pedestrian crossing. It is raining heavily, and you are getting wet because you do not have an umbrella. You need to run across the road to the shelter. However, you cannot immediately cross the road because a car is approaching. There is no other traffic on the road, but you have to wait until the car - which has the right of way - has passed. Yet, the driver of the car slows down, stops, and lets you cross the road first with a smiling face."

- Scenario 2: " Imagine you are a pedestrian and want to cross the street at a pedestrian crossing. You would have the right of way. But an approaching car honks, revs its engines, and accelerates, even though the driver clearly sees you, and prevents you from crossing the street. It is a very rainy day and the driver splashes you with a puddle of muddy water. As the car passes by, you notice that the driver and the passenger look at you mockingly as they enjoy the loud music in their car."
- Scenario 3: 'Imagine you are driving a car in a city center, stopping at a zebra crossing. As this street is crossed by many people, you have already spent some time waiting. A single pedestrian approaches the crossing when you finally thought you could continue driving. The pedestrian acknowledges that you have already waited, gives you their right of way, and steps back to allow you to pass."
- Scenario 4: "Imagine you are driving a car in a city center, stopping at a zebra crossing. As this street is crossed by many people, you have already spent some time waiting. A single pedestrian approaches the crossing when you finally thought you could continue driving. Although the pedestrian has seen you waiting for quite some time, they insist on their right of way and crosses. They walk extra slowly with the clear intention of annoying you."

B Definitions Provided in the Survey

Prosocial behavior *in traffic is defined as acting by taking the* ***well-being*** *of* ***other traffic participants*** *into account and* ***promoting effective cooperation*** *with others, such as drivers, passengers, pedestrians, and cyclists. Acting prosocial in traffic* ***benefits all traffic participants*** *in positive ways. It also helps to resolve traffic conflicts easily. This happens, for example, when searching for parking spaces, when merging lanes on the highway, when letting pedestrians cross the road, or thanking someone.* ***Aggressive or antisocial behavior*** *in traffic can be defined as opposite behavioral patterns, such as* ***acting and driving offensively, putting other people in danger, or acting selfishly in traffic****.*

In this part, we will provide you with some traffic scenarios. Please read each scenario from the beginning until the end, as some of the scenarios are similarly structured.



Please give your evaluations about the behavior of the other person as in the following example. If you think the behavior of the other person is good, then choose one of the points closer to good on the left side, depending on how good this behavior is for you.

Good ○●○○○○○ Bad

Similarly, if you think the behavior is rather bad, then choose a point closer to bad on the right side, depending on how bad you perceive it.

Good ○○○○○○● Bad

If you think the behavior was rather neutral, then choose a point closer to the middle.

Good ○○○●○○○ Bad

I think the other traffic participant / the behavior of other traffic participant was

Cooperative ○○○○○○○ Competitive

Helpful ○○○○○○○ Unhelpful

Courteous ○○○○○○○ Impolite

Prosocial ○○○○○○○ Antisocial

Trustworthy ○○○○○○○ Untrustworthy

Mindless ○○○○○○○ Mindful

Ungrateful ○○○○○○○ Thankful

Aware ○○○○○○○ Unaware

Inattentive ○○○○○○○ Attentive

Acknowledging ○○○○○○○ Ignoring

Unexpected ○○○○○○○ Expected

Predictable ○○○○○○○ Unpredictable

Fig. 5. Presentation of SPAT

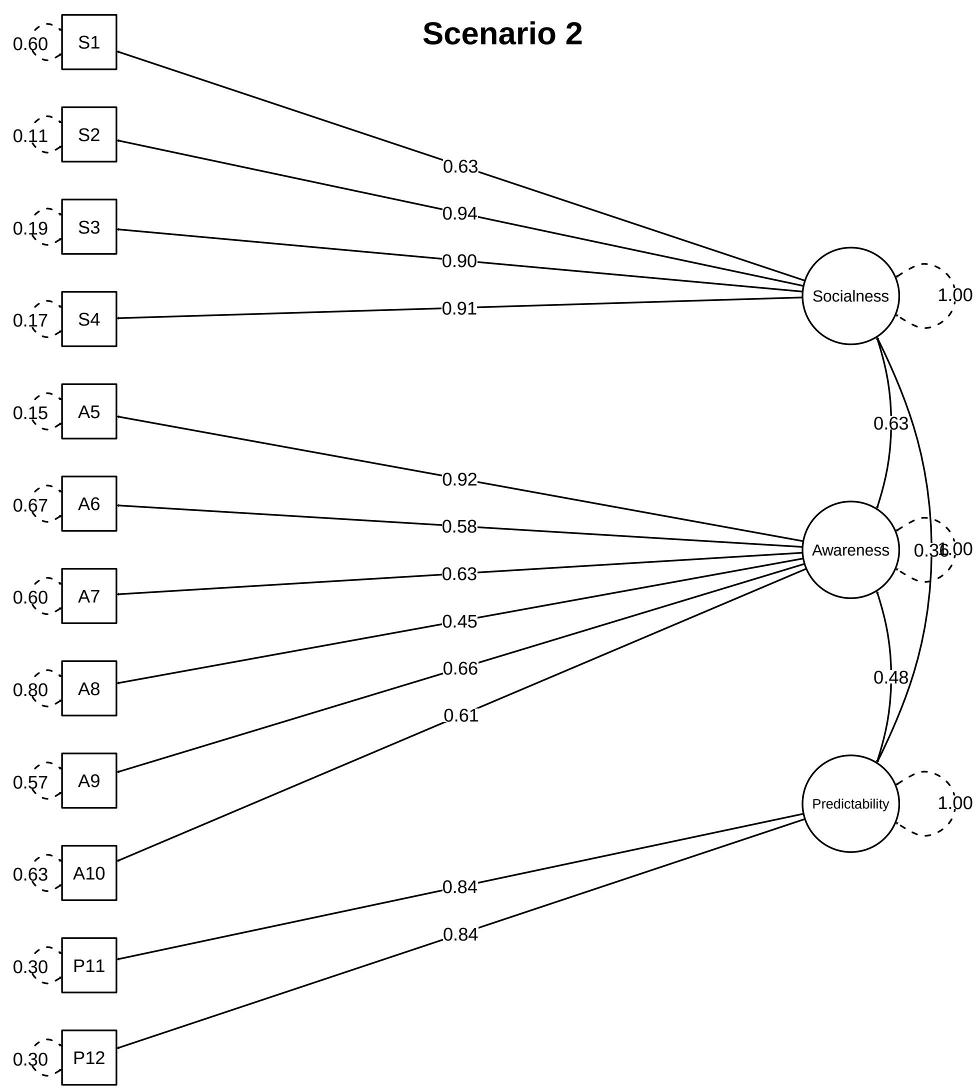


Fig. 6. Path diagram of Scenario 2. Participant from the pedestrian's perspective encounters an aggressive driver. Items are presented as squares with their associated error variances and latent variables as circles with their associated variances and covariance (edge). The values on the arrows are the loadings.

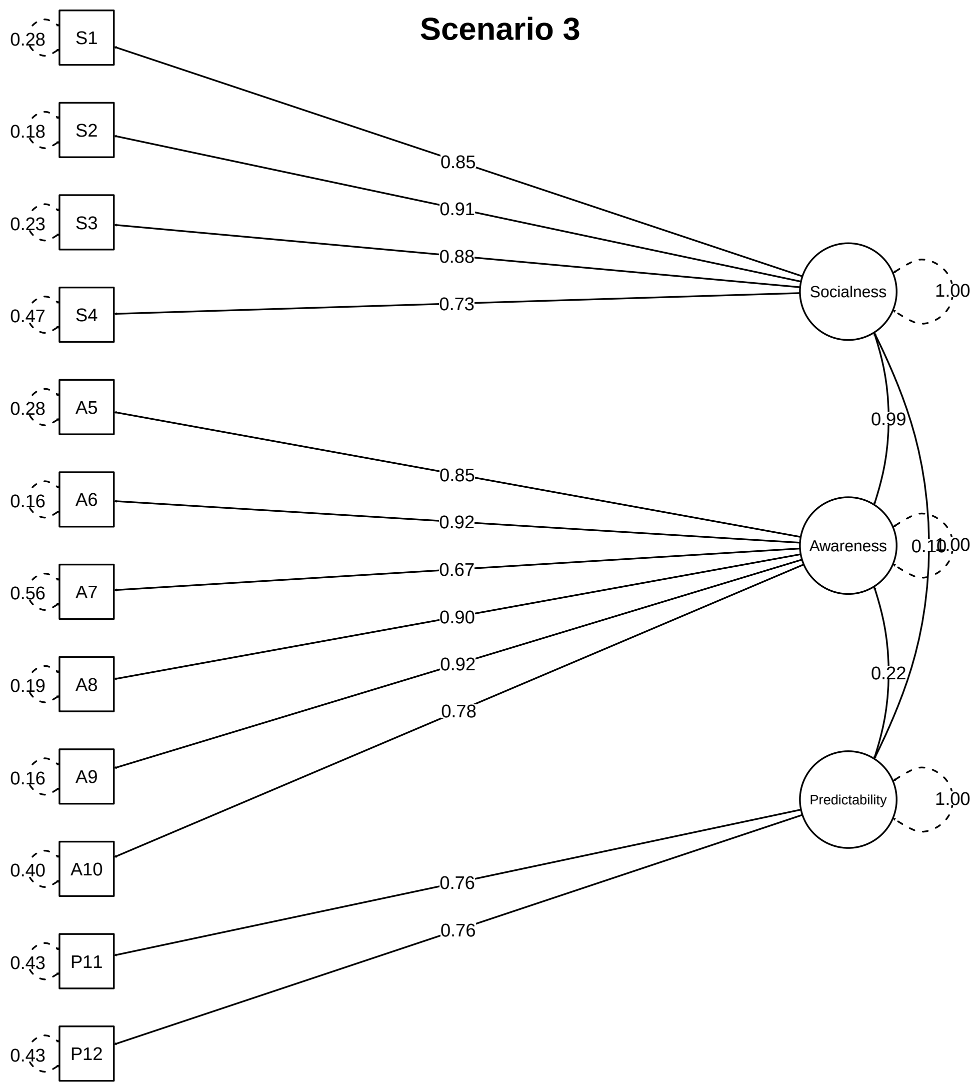


Fig. 7. Path diagram of Scenario 3. Participant from the driver's perspective encounters a prosocial pedestrian. Items are presented as squares with their associated error variances and latent variables as circles with their associated variances and covariance (edge). The values on the arrows are the loadings.

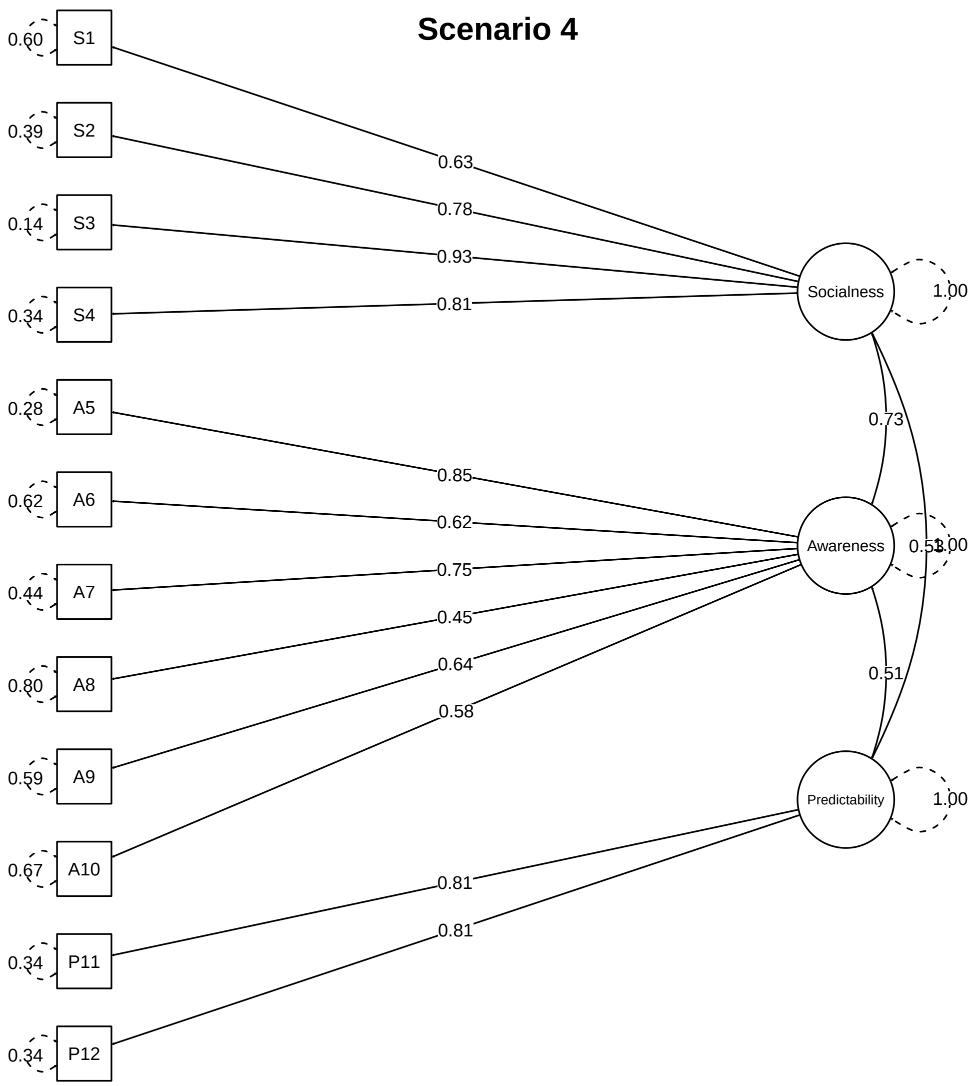


Fig. 8. Path diagram of Scenario 4. Participant from the driver's perspective encounters an aggressive pedestrian. Items are presented as squares with their associated error variances and latent variables as circles with their associated variances and covariance (edge). The values on the arrows are the loadings.

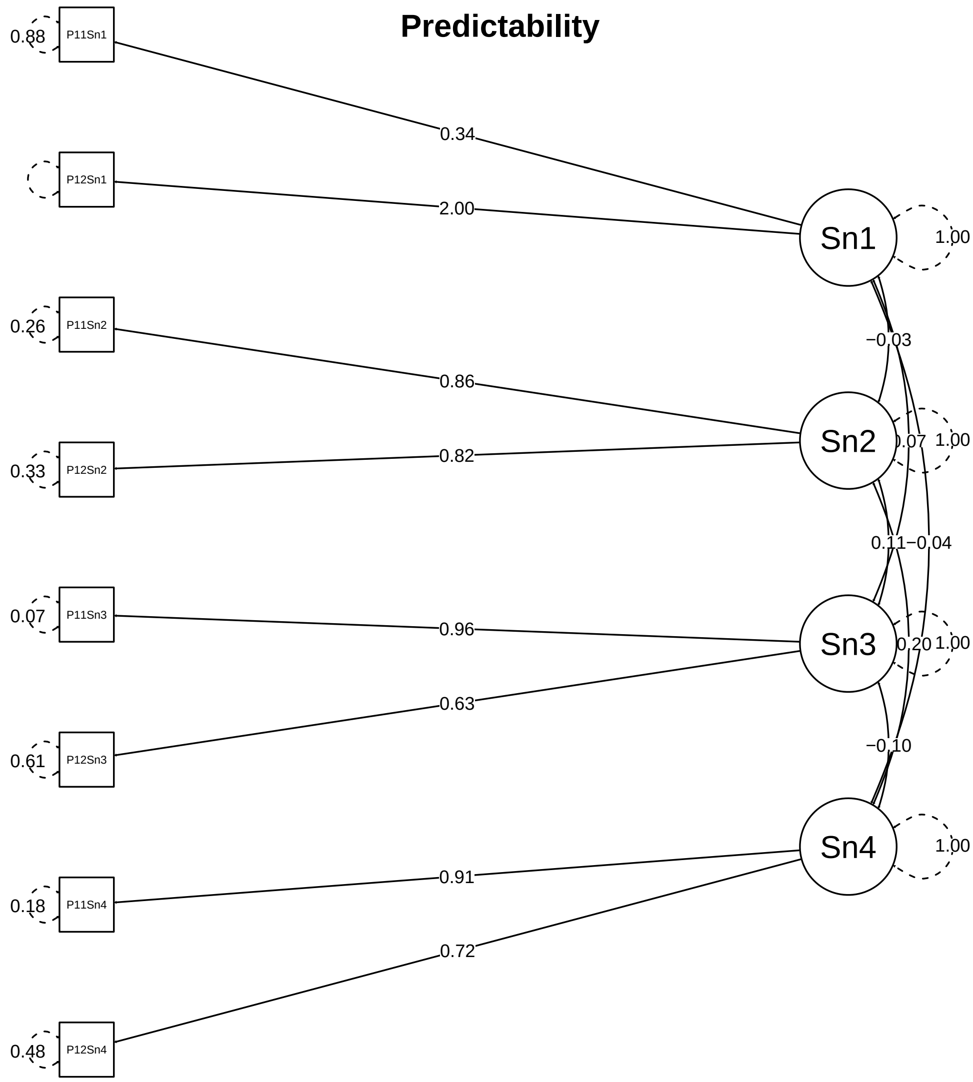


Fig. 9. Path diagram of factor Predictability and correlations of each scenario. Items are presented as squares with their associated error variances and latent variables as circles with their associated variances and covariance (edge). The values on the arrows are the loadings. Notation example: P11Sn1 indicates item P11 from Scenario 1.

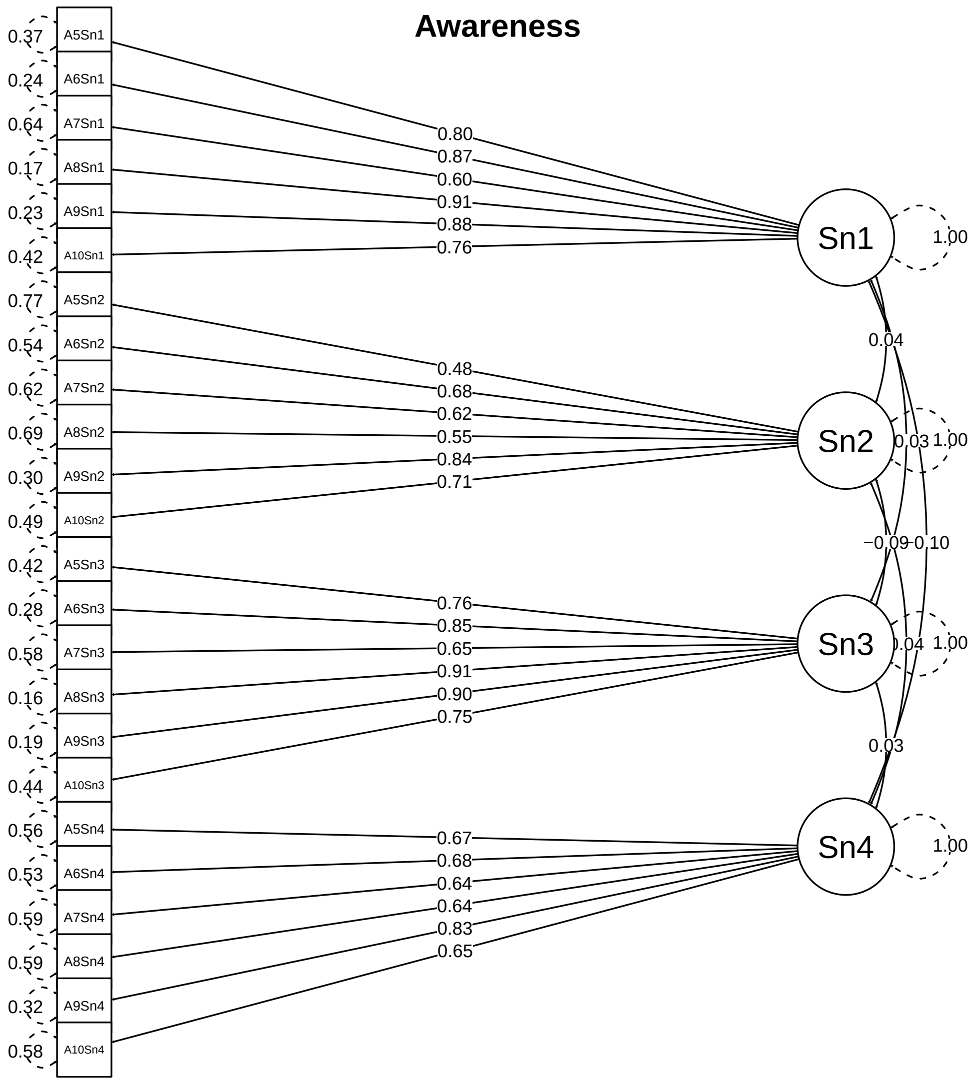


Fig. 10. Path diagram of factor Awareness and correlations of each scenario. Items are presented as squares with their associated error variances and latent variables as circles with their associated variances and covariance (edge). The values on the arrows are the loadings. Notation example: A5Sn1 indicates item A5 from Scenario 1.