A Face Validation Study for the Investigation of Proteus Effects Targeting Driving Behavior

Corinna A. Faust-Christmann[1(✉)], René Reinhard[2,3],
Alexandra Hoffmann[1,3], Thomas Lachmann[3,4], and Gabriele Bleser[1]

[1] Junior Research Group wearHEALTH,
Technische Universität Kaiserslautern, Kaiserslautern, Germany
christmann@cs.uni-kl.de
[2] Fraunhofer ITWM, Kaiserslautern, Germany
[3] Cognitive and Developmental Psychology, Center for Cognitive Science,
Technische Universität Kaiserslautern, Kaiserslautern, Germany
[4] Facultad de Lenguas y Educación, Universidad Nebrija, Madrid, Spain

Abstract. The Proteus effect describes the phenomenon that Immersive Virtual Reality users derive identity cues from their avatar's appearance which in turn activate specific stereotypes that influence the users' behavior or attitudes. The aim of this study was to validate faces of different age groups that can be used for avatars in further studies of Proteus effects targeting driving behavior. To achieve this goal, four neutral faces were selected from the CAL/PAL face database. They were rated with established questionnaires for driving behavior and driving styles to reveal explicit driver stereotypes. Implicit stereotypes were assessed with the implicit association test.

Study 1 (N = 93) revealed explicit driver stereotypes. In line with prior research, the young man's driving style was seen as riskier, angrier, with a higher velocity, and as less careful and less patient. The opposite pattern of results was found for the old woman, who received high scores for dissociative, anxious, and patient driving styles.

Study 2 (N = 160) replicated the overall pattern of results from study 1. Moreover, there were in-group out-group effects for the implicit gender stereotypes with regard to dissociative driving style and overall driving ability.

Based on these results, observable influences upon driving behavior are expected for driving errors, violations and lapses, as well as upon driving velocity with more violations and higher velocity for the male avatars and with more errors and lapses for the old female avatar. All faces were remodeled with Autodesk 3ds Max for further experiments investigating the Proteus effect targeting driving behavior.

Keywords: Explicit stereotypes · Implicit stereotypes · DBQ · MSDI · Avatar

1 Introduction

1.1 Proteus Effects

The Proteus effect describes the phenomenon that Immersive Virtual Reality (IVR) users derive identity cues from their avatar's appearance, e.g., height or age,

which in turn activate specific stereotypes that influence the users' behavior or attitudes [1]. There are first indications that these effects can persist for a short time after leaving IVR [2]. Although driving behavior is known to be associated with strong gender and age stereotypes (see Sect. 1.2 for details), possible Proteus effects have not yet been investigated in this context. This is rather surprising as there are first indications from simulator studies that driving behavior can be modulated by activating age [3] and gender stereotypes [4]. If the choice of a specific avatar actually influences subsequent driving behavior, special attention should be given to those avatars that trigger careful, attentive and skillful driving behavior to reduce the frequency of potentially dangerous driving errors.

1.2   Driver Stereotypes Regarding Age and Gender

Generally, two types of stereotypes can be distinguished. Explicit stereotypes refer to conscious thoughts and beliefs about a certain social group, whereas implicit stereo-types refer to unconscious attitudes [5].

Explicit driver stereotypes regarding age were so far only addressed in a few studies. Here, older drivers were rated as less aggressive than younger ones [6] and they were also perceived to have a lower likelihood of getting involved in accidents compared to younger drivers [7]. Moreover, one previous simulator study showed that priming an elderly stereotype by completing a scrambled-sentence task, resulted in lower maximum speed and longer driving time compared to a control condition [3]. In contrast to these findings, the only study on implicit driver stereotypes regarding age found "old" to be associated with dangerous drivers, whereas "young" was stronger associated with safe drivers. This finding was apparent for young and old participants [8]. In this study, the terms safe and dangerous drivers referred to driving skill (capable, skilled, ability) and different driving styles (awake, aware, focused, inattentive, risky), making it hard to deduce a certain driving behavior for older drivers.

Regarding gender stereotypes, men are generally perceived as risk takers and as more aggressive compared to women [9] and therefore are believed to be at a higher crash risk [7], to not comply to traffic rules [10, 11], and to drive more aggressively and at higher speed [10, 11]. These negative stereotypes are particularly pronounced with regard to young men [7]. The impact of these stereotypes was also apparent in a driving simulator study with young men, where it was revealed that priming with masculine words resulted in increased speeding compared to feminine or neutral words [4].

In contrast, the feminine stereotype is assumed to be passive, non-competitive and careful [11]. Following this idea, women are expected to drive carefully, at low speeds and to comply with traffic rules [10]. Moreover, women were more readily described as nervous drivers [12].

Despite the assumed risky and non-compliant driving style, men are, at the same time, believed to be more skilled drivers than women [7, 10, 13]. This shows that the difference between driving and safety skills is apparent in people's perceptions [14]. The former, driving skill, refers to driving performance, whereas the latter, safety skill, is determined by the driving style which someone chooses, for example a careful, patient or risky driving style. The perception of the two skills may, however, be linked.

An expected higher driving skill level may allow drivers to take more risks, while a fast and aggressive driving style may be considered a proof of driving skill [10].

1.3 Assessment of Driver Stereotypes

To allow a quantifiable approach for the assessment of explicit driver stereotypes regarding age and gender, pre-selected faces (one young man, one young woman, one old man, and one old woman, see Sect. 2.1 for details) were rated with established questionnaires for the assessment of driving behavior [15] and driving styles [16].

In this context, driving behavior usually refers to acts of aberrant behaviors while the driver is in control of a car. In the Driving Behaviour Questionnaire (DBQ) [15] safety-relevant driving behaviors are classified into three types: lapses, errors and violations. Lapses are absent-minded behaviors that do not pose any threat to others, whereas errors and violations can both be hazardous to others, but only the latter involve deliberate contraventions of traffic rules [17].

Driving style refers to how a person habitually drives, including choice of driving speed and level of attentiveness [18]. The following driving styles can be assessed with the Multidimensional Driving Style Inventory (MDSI) [16]: risky, angry, high velocity, dissociative, anxious, distress-reduction, patient, careful.

In addition to explicit stereotypes, implicit driving-related stereotypes can be assessed with the implicit association test (IAT), for example with regard to the driver's gender [8, 19]. An IAT is a reaction time based task that measures the strength and direction of an association between two dimensions (e.g. male/female and skilled/unskilled driver). In order to increase comparability between measures of explicit and implicit stereotypes, there was one IAT for driving behavior (skilled/unskilled) and several IATs for the driving styles: attentive/dissociative for the dissociative driving style, fast/slow for the high-velocity driving style, defensive/aggressive covering the patient, careful, angry and risky driving styles, as well as relaxed/distressed representing the anxious and stress reduction driving style.

Hypotheses on Driver Stereotypes
Based on the review of literature (see Sect. 1.2 for details), the following pattern of results are expected for the ratings of the four faces (see Sect. 2.1 for details) with MDSI [16] and DBQ [15]:

1. The young man receives higher scores for the risky, angry, and high velocity driving styles compared to the other three faces.
2. The women, especially the old woman, receive higher ratings for the anxious, distress-reduction, patient, and careful driving styles compared to the men.
3. Due to the assumed differences in driving skill, the women receive higher scores for lapses and driving errors.
4. The young man receives the highest scores for violations.

In addition, implicit driver stereotypes were investigated on an explorative basis, as the current literature does not afford clear expectations.

2 Methods

2.1 Face Selection

The aim of this study was to validate faces of different age groups that can be used in the process of avatar generation, which will become the basis for future studies of Proteus effects targeting driving behavior.

To achieve this goal, four neutral Caucasian faces (one young woman, one young man, one old woman, and one old man) were selected from the CAL/PAL face databank [20]. Care was taken to choose faces that only differ in gender and age, while controlling for further properties. Previous ratings reported by Ebner and colleagues indicated these faces to be comparable regarding overall attractiveness, likeability, distinctiveness, energy and mood [21]. In detail, the four faces were selected based on the following criteria:

1. The young faces of either gender were rated as younger than 31 years.
2. The old faces of either gender were rated older than 60 years.
3. The portrayed mood was rated as neutral by at least 60% of the participants.
4. All persons were rated as similar as possible with regard to attractiveness (means: 1.77–1.85), likeability (means: 1.69–1.92) and energy (means: 1.87–2.04) on a 0–4 scale [21].

2.2 Study 1: Explicit Age and Gender Stereotypes

The complete sample consisted of 93 adults. This sample consisted of both young adults (23 males, 22 females), at a mean age of 24.00 years (SD = 2.32; range 19–30 years), and an older sample (28 males, 20 females), at a mean age of 67.13 years (SD = 6.81; range 60–83 years). The sample of young adults consisted of students, while the participants in the older group consisted of retirees (62.5%), people still in the work process (27.1%), homemakers (8.3%) and unemployed persons (2.1%).

Participants rated the four selected faces using a German version of the DBQ [15] and the MDSI [16] to reveal explicit age and gender stereotypes for driving behavior and driving styles. The sequence of faces was counterbalanced between participants.

2.3 Study 2: Explicit and Implicit Gender Stereotypes

The sample consisted of 160 adults (75 males) with a mean age of 28.2 years (SD = 9.56, range: 18–65 years). Most participants were university students (58.75%) or employed (32.5%). The rest were unemployed (5.6%), retirees (1.9%) or homemakers (1.3%).

Participants had to rate two of the four faces (young man and young woman or old man and old woman) on a German version of the DBQ [15] and the MDSI [16] to reveal explicit gender stereotypes. The sequence of faces was counterbalanced between participants.

Moreover, they performed an IAT to reveal implicit stereotypes. The first dimension was always male/female. Each participant was randomly assigned to one of the following driving related dimensions: attentive/dissociative, skilled/unskilled, relaxed/distressed, fast/slow, or defensive/aggressive.

Based on a pilot test (N = 27 German university students) regarding the proximity of words to the target categories, the following German synonyms were chosen for the driving dimensions:

aufmerksam (attentive): achtsam, fokussiert, konzentriert, wachsam
nachlässig (dissociative): achtlos, abgelenkt, verzettelt, unbedacht
begabt (skilled): gut, geübt, talentiert, kompetent
unbegabt (unskilled): schlecht, hilflos, leistungsschwach, unfähig
gestresst (stressed): angespannt, erschrocken, nervös, ruhelos
entspannt (relaxed): gelassen, unverkrampft, locker, ruhevoll
schnell (fast): fix, rasend, eilig, zügig
langsam (slow): lahm, schleichend, bummelnd, trödelnd
defensiv (defensive): besonnen, ungefährlich, vorsichtig, zurückhaltend
aggressiv (aggressive): bedrohlich, ungezügelt, waghalsig, gefährdend

The seven-block version of the IAT was used [8, 22]. The first two blocks (24 trials each) were practice blocks with the category headings male and female in the first block and the driving related categories in the second block (e.g. skilled driver and unskilled driver). In the third (24 trials) and fourth block (40 trials) both category headings were combined (e.g. male and skilled driver on the left side and female and unskilled driver on the right side). The fifth block (40 trials) consisted of practice trials with the driving related category only, but the position of the headings was changed for left and right as compared to blocks 2, 3, and 4. In blocks 6 (24 trials) and 7 (40 trials) both category headings were again presented in combination (e.g. male and unskilled driver on the left side and female and skilled driver on the right side). Thus, data from blocks 3, 4, 6, and 7 were used in later analyses.

Participants sorted the target words as belonging to the respective driving related category and faces as belonging to the category male or female. Category pairings were displayed in the upper left and right corners of the computer screen. Words (8 for each category) and photographs (the same 4 as used in study 1) appeared in the middle of the screen in random order. Participants sorted them according to the correct category label by pressing a key on the keyboard that corresponded to the spatial location of the correct category.

The stimuli pair order was counterbalanced across participants. Thus, half of the participants would for example start in blocks 3 and 4 by sorting stimuli according to the category pairing male/skilled driver and female/unskilled driver, while the other half would start with the pairings of male/unskilled driver and female/skilled driver. They were then presented with the alternate pairing during blocks 6 and 7. The target stimuli remained on the screen until a response was recorded. Afterwards, feedback was displayed during 500 ms interstimulus intervals. Following trials with correct responses no stimulus was displayed in the center of screen, while a centrally displayed error symbol (X) followed incorrect responses.

IAT D scores were calculated by using the improved IAT scoring algorithm [22, 23]. The D score represents the difference in mean reaction time between the critical conditions divided by the standard deviations across conditions. Reaction times slower than 10,000 ms and faster than 300 ms were removed from the data set prior to D score calculations. Error trials were not removed from the analysis in accordance with Greenwald, Nosek and Banaji [23].

3   Results

3.1   Study 1

The data obtained from each of the first study's questionnaires (MDSI and DBQ) were analyzed using a mixed-design multivariate analysis of variances (MANOVA) with age group and gender of the participant as between subject factors and age and gender groups of the photograph as within subject factors. For the MDSI data the MANOVA revealed a main effect of participant's age group, Wilk's λ = .62, $F(8,82)$ = 6.28, $p <$ .01, a main effect of photograph's age group, Wilk's λ = .35, $F(8,82)$ = 18.87, $p < .01$, as well as a main effect of gender group of photograph, Wilk's λ = .37, $F(8,82)$ = 17.45, $p < .01$. Moreover, there was an interaction between the participant's age group and the age group of the photograph, Wilk's λ = .66, $F(8,82)$ = 5.32, $p < .01$, between participant's age group and the photograph's gender group, Wilk's λ = .83, $F(8,82)$ = 2.16, $p = .04$, and lastly between the age group and gender group of the photographs, Wilk's λ = .56, $F(8,82)$ = 8.07, $p < .01$.

Comparable results were found for the DBQ scales. Here, a main effect of participant's age group, Wilk's λ = .78, $F(3,87)$ = 8.04, $p < .01$, a main effect of the photograph's age group, Wilk's λ = .47, $F(3,87)$ = 32.91, $p < .01$, and a main effect of the photograph's gender group, Wilk's λ = .36, $F(3,87)$ = 51.67, $p < .01$ were found. Additionally, the interactions between the participant's and the photograph's age group, Wilk's λ = .79, $F(3,87)$ = 7.60, $p < .01$, between the participant's age group and the photograph's gender group, Wilk's λ = .89, $F(3,87)$ = 3.62, $p = .02$, and between the photograph's age and gender group, Wilk's λ = .72, $F(3,87)$ = 11.05, $p < .01$. All remaining main effects and interactions did not reach significance.

To further explain this pattern of results, additional univariate analyses of variance (ANOVA) were conducted for each scale of the MDSI and DBQ. Alpha level was adjusted according to the Holm-Bonferroni method to consider effects of multiple testing. The main effect of the participant's age group was found for all scales (all $p <$ .01, except the MDSI's anxious driving, $p = .02$, and the DBQ's violations scale, $p =$ .03) with the exception of risky ($p = .14$), and patient driving ($p = .24$). The main effect of the photograph's age group was present in all scales (all $p < .01$) apart from distress reduction driving ($p = .32$). At the same time, the main effect of the photograph's gender group was apparent in each scale (all $p \leq .02$). The interaction between the participant's and the photograph's age group was found for dissociative, risky, and patient driving, as well as for errors (all $p < .01$). The interaction between the photograph's age and gender group was apparent for dissociative driving, risky, distress reduction, and careful driving, as well as for errors and lapses (all $p < .01$).

All remaining main effects and interactions did not reach significance for any of the scales. Ratings for the four photographs were compared with post-hoc t-tests, separately for the young and old subgroup. The results for each MDSI scale, including post-hoc comparisons between photographs, are pictured in Fig. 1 for the young subsample and in Fig. 2 for the old subsample. The results of the DBQ for both the young and old groups are displayed in Fig. 3.
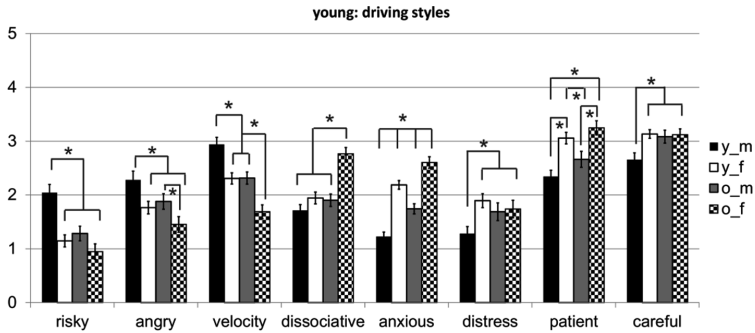
Fig. 1. Mean ratings from the young group of participants for the four faces (young male = y_m, young female = y_f, old male = o_m, old female = o_f) regarding the eight driving styles based on the MDSI. Error bars indicate standard errors of the mean.
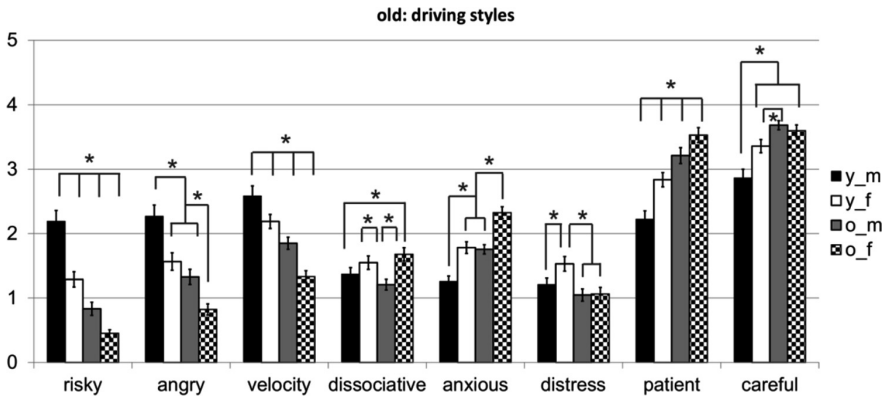
Fig. 2. Mean ratings from the old group of participants for the four faces (young male = y_m, young female = y_f, old male = o_m, old female = o_f) regarding the eight driving styles based on the MDSI. Error bars indicate standard errors of the mean.
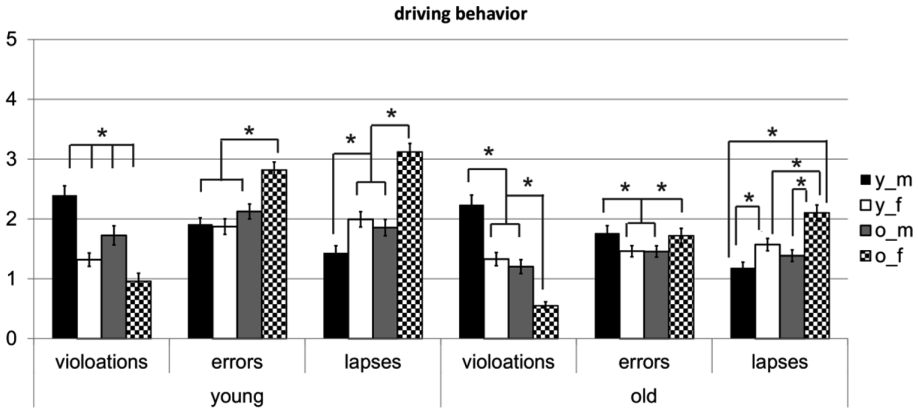
Fig. 3. Mean ratings from the young group of participants (left side) and the old group of participants (right side) for the four faces (young male = y_m, young female = y_f, old male = o_m, old female = o_f) regarding driving behavior based on the DBQ. Error bars indicate standard errors of the mean.

3.2   Study 2

Comparable with Experiment 1, a MANOVA was conducted for each questionnaire data set (MDSI and DBQ). The age group of the photograph had been introduced as a between subject factor to reduce overall testing time. For the MDSI data the MANOVA revealed a main effect of the participant's gender, Wilk's $\lambda = .89$, $F(8,149) = 2.40$, $p = .02$, a main effect of the photograph's age group, Wilk's $\lambda = .72$, $F(8,149) = 7.23$, $p < .01$, a main effect of the photograph's gender group, Wilk's $\lambda = .51$, $F(8,149) = 18.25$, $p < .01$ and an interaction between the photograph's age and gender group, Wilk's $\lambda = .79$, $F(8,149) = 4.94$, $p = < .01$. For the DBQ scales, there was a main effect of the age group of photograph, Wilk's $\lambda = .71$, $F(3,154) = 21.00$, $p < .01$, as well as a main effect of the gender group of photograph, Wilk's $\lambda = .49$, $F(3,154) = 53.21$, $p < .01$. The interaction between the participant's and the photograph's age group also reached significance, Wilk's $\lambda = .84$, $F(3,154) = 9.51$, $p < .01$.

This pattern of results was further explored using additional univariate analyses of variance (ANOVA) for each scale of the MDSI and DBQ with Holm-Bonferroni alpha level adjustment.

The main effect of the photograph's age group was present for angry, high velocity, anxious, and careful driving, as well as for all scales of the DBQ scales (all $p < .01$). The main effect of the gender group of the photograph was found for all scales (all $p < .01$) with the exception of risky ($p = .92$) and distress reduction driving ($p = .74$). The interaction between age and gender group of the photograph was apparent for dissociative, risky, and careful driving, as well as for errors and lapses (all $p < .01$). The mean ratings for each MDSI scale are shown in Fig. 4 and for each DBQ scale in Fig. 5.
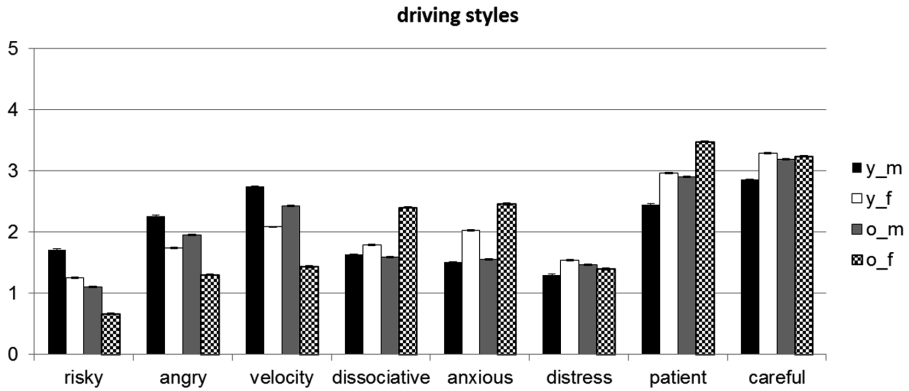
driving styles

Fig. 4. Mean ratings for the four faces (young man = y_m, young woman = y_f, old man = o_m, old woman = o_f) regarding the eight driving styles based on the MDSI. Error bars indicate standard errors of the mean.
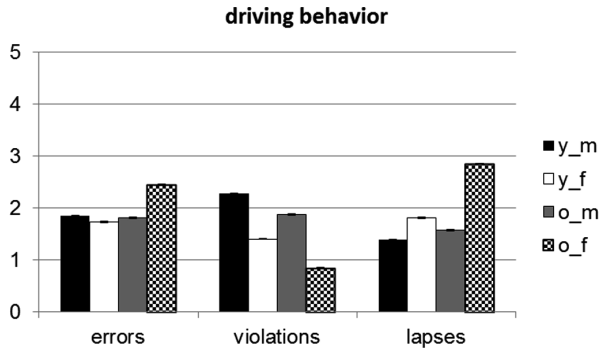
driving behavior

Fig. 5. Mean ratings for the four faces (young man = y_m, young woman = y_f, old man = o_m, old woman = o_f) regarding driving behavior based on the DBQ. Error bars indicate standard errors of the mean.

Significant group differences in D scores were found for the attentive/dissociative category, $t(33) = -4.57$, $p > .001$, and the skilled/unskilled category, $t(31) = -4.67$, $p > .001$, with higher D scores for female than for male participants (see Fig. 6), indicating that female participants more strongly associated attentive and skilled with female compared to male participants, who more strongly associated attentive and skilled with male.
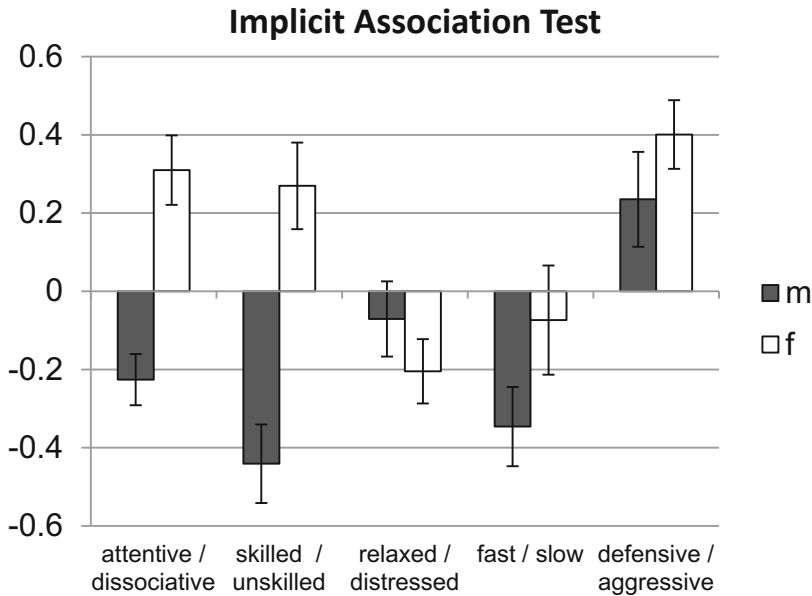
Fig. 6. Mean D scores for male (m, grey boxes) and female participants (f, white boxes). D scores above zero indicate stronger associations between the female attentive/skilled/relaxed/fast/defensive and the male-dissociative/unskilled/distressed/slow/aggressive category pairings. Error bars indicate standard errors of the mean.

4 Discussion

The goal of this study was to evaluate if pre-selected faces (see Sect. 2.1 for details) are associated with distinctive driver stereotypes related to their perceived age and gender. Study 1 dealt with explicit stereotypes on driving behavior and driving styles. Ratings were made by young and old adults. Study 2 replicated the explicit ratings with a larger sample and also included a set of IATs to reveal implicit stereotypes on driving behavior and driving styles.

4.1 Study 1

In line with the expectation that young men are perceived as risk takers (hypothesis 1), the photograph of the young man received the highest scores for the risky, angry, and high velocity driving styles. This finding was evident both in the ratings of young and old participants. Moreover, in line with the stereotype of the passive, non-competitive female driver (hypothesis 2), the women received higher ratings for the patient, and careful driving styles compared to the young man. Additionally, comparable with the previous attribution of females as nervous drivers [12], the driving style of women was described as more anxious and geared towards distress-reduction. In line with prior findings that old drivers are expected to be less aggressive [6] and to be at lower risk for accidents [7], the old man also received high ratings for the careful driving style.

In compliance with stereotypes of higher driving skills in male drivers (hypothesis 3, also see [7, 10, 13]), there was a main effect of the gender group of the photograph on ascribed lapses and errors. This pattern was particularly pronounced for ratings of the old woman. In line with hypothesis 4, the opposite pattern of results was found for violations. Here, the young man received the highest score, followed by the young woman and the old man, while the old woman received the lowest score.

By comparing the ratings of the young and old subsample two main findings become apparent. First, that the overall pattern of results is quite similar in both subsamples. The results underpin the existence of explicit stereotypes of young men as risky, angry and high velocity drivers committing a lot of violations and of especially old women as patient and anxious drivers with a lot of lapses and errors. Second, by comparing the ratings of the young and old subsample for the old man, pronounced in-group out-group effects become apparent. For the young subsample, the only difference between the young woman and the old man was found for the patient driving style (higher ratings for the young woman). By contrast, the old subsample rated the old man's driving style as less risky and dissociative, as prone to lower velocities, and to be more patient and careful compared to the young woman. The group serving bias is also apparent for the DBQ ratings. Here, the old woman received much lower scores by the old subsample compared to the young subsample. This is the first demonstration of age group specific biases. No in-group out-group effects were observed with regards to gender. This is inconsistent with previous reports of gender-specific stereotypes. Here, it had been reported that females rated men's likelihood of accidents as a higher than male raters [7].

4.2 Study 2

Explicit Stereotypes
With regard to explicit stereotypes, the second study was able to replicate the patterns observed in the first study. In line with the first study, the young man received the highest scores for the risky, angry, and high velocity driving style, whereas the old woman received the highest scores for dissociative, anxious and patient driving style. Moreover, for the pattern of ascribed errors, violations and lapses were replicated, with most errors and lapses attributed to the old woman and most violations to the young man. The second study, could again not establish any in-group out-group effects with regard to gender.

Implicit Stereotypes
In-group out-group effects with regard to gender were, however, apparent in the implicit stereotype measurements. Women were strongly associated with attentive and skilled drivers, whereas the opposite pattern of results was found for men. For defensive and slow drivers, both male and female raters showed a tendency to associate defensive drivers with female and aggressive drivers with male. The implicit association of male drivers with an aggressive and female drivers with a defensive driving style is consistent with the findings from the observed explicit driving stereotypes.

The current study is the first to reveal implicit driver stereotypes with respect to gender (but see [8] for implicit driver stereotypes regarding age). Further studies are, however, needed to substantiate the findings of implicit driver stereotypes.

4.3 Outlook

To summarize, the four faces are associated with explicit and implicit driver stereotypes. For experiments on the Proteus effect, observable influences upon driving behavior are expected for driving errors, violations and lapses, as well as upon driving velocity with more violations and higher velocity for the young male avatar and with more errors and lapses for the old female avatar.

All faces were remodeled with Autodesk 3ds Max for further experiments investigating the Proteus effect regarding walking speed [24] and driving behavior based on these results (see Fig. 7). Body models for each gender and age group were constructed based on data from a representative serial measurement campaign of German adults conducted between 2007 and 2008 by the Forschungsinstitut Hohenstein Prof. Dr. Jürgen Mecheels GmbH & Co.KG and the Human Solutions GmbH.

Fig. 7. Avatar faces based on the four faces.

An additional questionnaire based study (N = 50, 22−81 years) asked for ratings of both the original photos of the faces and the resulting avatars with regard to their attractiveness, likeability, energy and perceived age [21] to ensure comparability of the avatars with the original faces. Spearman's rank correlation analysis revealed systematic associations between the face and avatar ratings. The correlations were significant for all combinations of the same face and avatar (all $p \leq .05$), for the old man

(attractiveness: ρ = .341, likeability: ρ = .401, energy: ρ = .423, perceived age: ρ = .396), the old woman (attractiveness: ρ = .565, likeability: ρ = .343, energy: ρ = .520, perceived age: ρ = .424), the young man (attractiveness: ρ = .499, likeability: ρ = .555, energy: ρ = .493, perceived age: ρ = .414), and the young woman (attractiveness: ρ = .324, likeability: ρ = .334, energy: ρ = .337, perceived age: ρ = .457).

In a further study, 67 young adults (18–34 years) either experienced the young (22 participants) or old avatars of their own gender (23 participants) in IVR, or did not enter IVR (22 participants). They then rated the two avatars of their own gender with regard to their anticipated walking speed. One-sample upper-tailed z-tests for dichotomous outcomes revealed that the elderly avatars were rated as slower more often than would be expected by chance. This was apparent for participants who had previously embodied the older avatar, z = 2.294, p = .022, in the ratings of participants who had embodied the young avatar, z = 1.705, p = .044, and for the control group who had not entered IVR, z = 2.558, p = .016.

The avatar's different age groups were further successful in eliciting Proteus effects on real life walking speed after participants had left IVR, as participants tended to traverse a set distance slower after embodying the older gender-matched avatars than either the group of participants who had previously embodied the younger avatars or a non-IVR control group [24]. Future studies will explore, whether the avatars can elicit similar effects on objective driving behaviors in driving simulators, e.g., on the choice of driving speed. In these studies, the temporal stability of such Proteus effects after leaving IVR, as well as the preconditions for their occurrence, will be of particular interest.

References

1. Yee, N., Bailenson, J.: The proteus effect. The effect of transformed self-representation on behavior. Hum. Commun. Res. **33**, 271–290 (2007)
2. Yee, N., Bailenson, J.N., Ducheneaut, N.: The proteus effect. Commun. Res. **36**, 285–312 (2009)
3. Branaghan, R.J., Gray, R.: Nonconscious activation of an elderly stereotype and speed of driving. Percept. Mot. Skills **110**, 580–592 (2010)
4. Schmid Mast, M., Sieverding, M., Esslen, M., Graber, K., Jäncke, L.: Masculinity causes speeding in young men. Accid. Anal. Prev. **40**, 840–842 (2008)
5. Greenwald, A.G., Banaji, M.R.: Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychol. Rev. **102**, 4–27 (1995)
6. Davies, G.M., Patel, D.: The influence of car and driver stereotypes on attributions of vehicle speed, position on the road and culpability in a road accident scenario. Leg. Criminological Psychol. **10**, 45–62 (2005)

7. Glendon, A.I., Dorn, L., Davies, D.R., Matthews, G., Taylor, R.G.: Age and gender differences in perceived accident likelihood and driver competences. Risk Anal. **16**, 755–762 (1996)

8. Lambert, A.E., Seegmiller, J.K., Stefanucci, J.K., Watson, J.M.: On working memory capacity and implicit associations between advanced age and dangerous driving stereotypes. Appl. Cognit. Psychol. **27**, 306–313 (2013)

9. Locksley, A., Borgida, E., Brekke, N., Hepburn, C.: Sex stereotypes and social judgment. J. Pers. Soc. Psychol. **39**, 821–831 (1980)

10. Granié, M.-A., Papafava, E.: Gender stereotypes associated with vehicle driving among French preadolescents and adolescents. Transp. Res. Part F Traffic Psychol. Behav. **14**, 341–353 (2011)

11. Simon, F., Corbett, C.: Road traffic offending, stress, age, and accident history among male and female drivers. Ergonomics **39**, 757–780 (1996)

12. Berger, M.L.: Women drivers!: the emergence of folklore and stereotypic opinions concerning feminine automotive behavior. Women's Stud. Int. Forum **9**, 257–263 (1986)

13. DeJoy, D.M.: An examination of gender differences in traffic accident risk perception. Accid. Anal. Prev. **24**, 237–246 (1992)

14. Lajunen, T., Summala, H.: Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments. Pers. Individ. Differ. **19**, 307–318 (1995)

15. Reason, J., Manstead, A., Stradling, S., Baxter, J., Campbell, K.: Errors and violations on the roads: a real distinction? Ergonomics **33**, 1315–1332 (1990)

16. Taubman-Ben-Ari, O., Mikulincer, M., Gillath, O.: The multidimensional driving style inventory—scale construct and validation. Accid. Anal. Prev. **36**, 323–332 (2004)

17. Parker, D., Reason, J., Manstead, A., Stradling, S.: Driving errors, driving violations and accident involvement. TERG **38**, 1036–1048 (1995)

18. Taubman-Ben-Ari, O., Skvirsky, V.: The multidimensional driving style inventory a decade later: review of the literature and re-evaluation of the scale. Accid. Anal. Prev. **93**, 179–188 (2016)

19. Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K.: Measuring individual differences in implicit cognition: the implicit association test. J. Pers. Soc. Psychol. **74**, 1464–1480 (1998)

20. Minear, M., Park, D.C.: A lifespan database of adult facial stimuli. Behav. Res. Methods Instrum. Comput. **36**, 630–633 (2004). A Journal of the Psychonomic Society, Inc.

21. Ebner, N.C.: Age of face matters: age-group differences in ratings of young and old faces. Behav. Res. **40**, 130–136 (2008)

22. Lane, K.A., Banaji, M.R., Nosek, B.A., Greenwald, A.G.: Understanding and Using the Implicit Association Test: IV: What We Know (So Far) about the Method. Guilford Press, New York (2007)

23. Greenwald, A.G., Nosek, B.A., Banaji, M.R.: Understanding and using the implicit association test: I. an improved scoring algorithm. J. Pers. Soc. Psychol. **85**, 197–216 (2003)

24. Reinhard, R., Shah, K.G., Faust-Christmann, C.A., Lachmann, T.: Acting your avatar's age: effects of virtual reality avatar embodiment on real life walking speed. Media Psychology (in press)