1. INTRODUCTION

Human-machine interfaces (HMI) in complex systems and environments are becoming crucial touchpoints to engage with users as more products and services are digitised or automated. Humans are taken out of the loop and replaced by machines, underscoring the importance of well-designed HMI. While usability and user acceptance studies cover some aspects of user experience, experience-driven design focuses more on emotional and experiential elements. Experience-driven design allows for direct intervention with the user to shape their perceptions of experiences rather than attempting to disentangle entire ecosystems and solve wicked problems. One such example of complex systems can be found in public transport. The elements of a multi-stakeholder environment, service operations and infrastructure combine, sometimes resulting in conflicting goals. Challenges are amplified with the introduction of autonomous vehicle (AV) technology, which many countries are currently experimenting with. This includes Singapore, the site of this study, which aims to introduce AVs for both passenger and industrial solutions.

Qualitative investigations into commuting reveal mobility experiences to be much more complex than simply a behaviour. They are instead amalgamations of beliefs, values and activities, coupled with norms and intentions. One such dimension of the experience is emotional, reflecting the embodied nature of commuting, i.e., the actual bodily experience and how it engages with the environment during travel. Mobility should therefore focus not only on transporting people from origin to destination. Bissell describes the ‘affective atmosphere’ of public transport, referring to the propensity of a space in time to generate certain events, actions, feelings and emotions. The affective atmosphere is intangible, not entirely consciously experienced, and is shaped by fellow passengers, time of travel or unexpected service interruptions. It is here where interaction touchpoints can influence passenger experiences by provoking specific emotional responses. The overarching goal is thus to alter the psychological characteristics of mobility experiences rather than structural interventions, which require major changes to the external environment. In the case of AV transport systems, it is unclear how the sociality of the interface between human and machine should be designed. The various use cases in public transport scenarios further confound correlations of factors affecting user perceptions of their interactions and consequently, their travel experiences.

Therefore, establishing links between emotional response and the stimulus – comprised of the HMI and scenario – can lead to a more informed interaction design process within different autonomous mobility contexts to compel user emotions. In this study, we posit that people’s emotional responses can be used as a measure to evaluate different design concepts, leading to the identification of these links. We thus aim to discover the effects of HMI design concepts and mobility scenarios on emotional
responses and satisfaction, so that we can define guidelines for the design of interfaces in future mobility systems. This paper is based on the methodology and preliminary findings described in [7] and presents the full results from the completed study. The case study which the research takes as context is a 30-seater autonomous vehicle (AV) developed for public transport in Singapore. As the human driver is taken out of the loop, two possible HMI solutions were proposed in the form of a Virtual Companion and Robot Companion. Both seek to fulfill the customer care and information roles formerly played by human staff.

2. LITERATURE REVIEW

Emotion is subconsciously experienced, can be intuitively accessed and can be attributed to an artefact, i.e., a cause such as a thing, person or event [8]. It is thus a universal dimension of human factors suitable for studies across demographics and contexts. Empathic design tries to understand the user’s mental process and use that newfound meaning by “…gaining understanding of the user (cognitive) through the user’s emotional state (affective)” [9]. Rather than directly evaluating design concepts, the use of emotion to mediate the evaluation seeks to reveal more organic and accurate responses. Dunn & Schweitzer [10] demonstrated that emotions influence trust in two interesting ways. First, the incidental emotional states of subjects influenced their levels of trust, where happy or grateful participants were significantly more trusting than sad or angry participants. Secondly, when participants can identify the source of incidental emotion, the influence of emotion on trust was negated. Understanding how emotional manipulation can affect trust levels could be useful in shaping public perceptions of products or services, making such knowledge especially pertinent in discussions on future technology.

In addition to trust, emotions are used to guide our decision-making [11]. People seek out experiences they feel positively about. But maintaining positive emotion throughout entire experiences can be a challenge. It could also have the opposite effect, being detrimental to perceptions of experiences after the fact. Contrast in emotional journeys, i.e., an experience with both positive and negative emotion, has been found to achieve the intent of positive experiences more effectively than a more consistent, monotonous ride [12]. The peak-end rule states that the most extreme emotions and emotions felt at the end of experiences can have a greater influence on individual perceptions of past experiences [13]. This suggests room for the occurrence of both positive and negative emotion in a user experience while still creating an overall positive experience. Thus, by identifying how interventions can be designed for particular emotional responses at different touchpoints in the user journey, we will be able to craft user experiences to elicit desired emotional experiences.

Applications of emotional design to the mobility experience are limited. The kansei engineering method was extended to the design of a new bus interior by considering customer input to generate kansei words [14]. The method was successful in identifying experience requirements of the customer experience (e.g., shapes or gangway lighting) which impacted kansei words. We conducted a similar study to define design requirements [15] based on the Kano model of product innovation, where product attributes are classified into Must-meets (must-haves) and Delight (ideal-to-haves) categories, among others [16]. Must-haves are attributes which will cause dissatisfaction when absent, though their presence does not lead to additional satisfaction. Ideal-to-haves greatly improve satisfaction when present but are not detrimental when absent. Results contributed to stimulus design in this study and showed that context plays an important role. In the study [15], participants preferred human assistance over assistance from a Virtual Companion when feeling unwell as they preferred a human touch. Literature in this area is limited, so this study attempts to explore the effects of scenario types, among others, on preferred support types in future mobility. The personality of the participant has been shown to influence preferences for Companion types, where participants who are more introverted prefer interacting with less anthropomorphic robots [17]. The participant’s personality thus plays an important role in Companion design and interaction. Based on the bipolar competence-frustration model within the framework of appraisal theory, emotional responses are cognitive processes used to evaluate event outcomes [18]. Successful task completion leading to the achievement of user goals generates positive emotion, while negative emotion such as frustration and anxiety may occur with task completion failure. People need to feel that the designed HMI helps rather than hinders them from achieving their goals. These goals can be identified through user research on people’s mobility experiences. It should be noted that while people feel more positively about experiences when the need is fulfilled to a greater degree, attribution to a source moderates the effect [19]. Subjects need to be able to identify the cause of their emotion as well as distinguish between instrumental, pragmatic qualities and non-instrumental, hedonic qualities. Isolating factors to define cause and effect between function and scenario will thus be crucial for the investigation.
3. METHODOLOGY

3.1 Hypotheses

The first hypotheses (H1, H2 and H3) reflect direct relationships between subjects’ evaluations, i.e., emotional response and level of satisfaction, and the stimulus, which is composed of the scenario and Companion. As defined in a previous study by the authors [15], functions are split into two categories, must-haves and ideal-to-haves, which affect user satisfaction in different ways. They are hypothesised to generate the following emotional responses for a within-subjects analysis:

- H1a: the presence of must-haves will not cause positive emotion
- H1b: the absence of ideal-to-haves will not cause negative emotion
- H2: the absence of must-haves will cause negative emotion
- H3: the presence of ideal-to-haves will cause positive emotion

Secondly, we anticipate an effect of scenarios on subjective evaluations of function (H4) in an exploratory hypothesis.

The third relationship for analysis posits that participant traits will influence Companion preference in a between-subjects test:

- H5a: higher comfort with technology will lead to more positive evaluations of the Companions
- H5b: males will have more positive evaluations of the Companions than females
- H5c: higher introversion will lead to higher ratings for the less anthropomorphic Virtual Companion

3.2 Stimulus Design

The stimulus was built through a combination of visuals, background noise and a text-based narrative and showcased functions in different scenario contexts. The scenes (Figure 1) were shown from the point of view of the participant to immerse them in the scenarios as if they were the protagonist in a movie. In certain scenes, they were depicted from a third-person perspective and were rendered as a blue figure. The visuals were created using people silhouettes, 3D renders of the two Companions and background photographs. Scenes were placed in Microsoft PowerPoint and run automatically on a timer. Background noise simulated the environment for greater immersion in the narrative, e.g., traffic white noise, crowded spaces, buses driving. The narrative contained a short introduction, situational problem and solution (as provided by the Companion), all written from the perspective of the participant. Three scenarios were constructed to test different features of the Companion: a commuter scenario, emergency scenario and tourist/unfamiliar scenario. The scenarios were based on common use cases which occur on public transport in Singapore but take place in a future where autonomous buses are already in operation.

In the commuter scenario, the route is a known and regular one. A problem arises when, unbeknownst to the subject, the autonomous bus is delayed. The information required in this scenario relates to vehicle arrival times and traffic conditions. The emergency scenario takes place in the AV. A problem occurs when the other passenger experiences a medical issue. Finally, the tourist/unfamiliar scenario is set in an unfamiliar bus interchange with an unknown bus route. A crowded environment and time urgency combine to create the problem. The scenarios were developed to provide problems relating to two types of assistance: travel information (must-haves) and physical assistance (ideal-to-haves). These relate to H1, H2 and H3.
In each scenario, the Companion comes in two forms (Figure 2). The Virtual Companion is a non-biomimetic display screen which is more static (lacking initiative), informative (by providing pertinent travel information) and has more cumbersome input (touchscreen with digital buttons and an onscreen keyboard). It represents must-have functions. The Robot Companion is an anthropomorphic, mobile robot which is more proactive (approaches the user without prompting), not informative (instead, providing physical assistance) and can be interacted with verbally. It represents ideal-to-have functions. The assignment of these two types of functions was based on previous work [15], where we found that participants rated practical, trip-related information as core features, i.e., must-haves, and non-trip related functions such as personalisation and physical assistance as bonus, i.e., ideal-to-have features.

3.3 Study Design and Procedure

A questionnaire collected subjective evaluations and background information. The two dimensions of the Circumplex Model of Affect [8] were used to measure subjects’ emotional responses because the use of underlying constructs break down people’s emotions in a measurable way while not depending on individual abilities to articulate emotion [11]. Using a continuum was also preferred to overcome sociocultural and linguistic differences which may arise with the use of discrete categories such as happy, sad, or frustrated, reduce selection bias on the part of the researcher, and avoid priming of participants. Three Likert scales with eleven anchor points (middle point was neutral) measured valence (pleasant-unpleasant), arousal (excited-bored) and satisfaction with the Companion. After experiencing both Companions in each scenario, participants were asked to select the Companion they preferred for that scenario and explain why. They also rated each Companion on traits such as trustworthiness, comfort and approachability. Participant personality was measured using the BFI-10 through items like extraversion and neuroticism [20]. Comfort with technology was measured based on an adaptation of the Personal Innovativeness in Information Technologies scale [21]. Other demographic questions included age, gender, and education level.

Due to the COVID-19 pandemic in Singapore, two study conditions were used for data collection to adapt to the restrictions put in place. The original condition was an on-site study at the researchers’ offices, while the adapted condition was deployed online. In the on-site condition, participants were assigned to one of three stations consisting of a laptop, mouse, on-ear headphones and the paper-based questionnaire. The stations were set up in a quiet room so participants could begin the study as they arrived. Starting times were deliberately staggered to minimize exposing subjects to time pressure. Online data collection was conducted via video conferencing using Zoom [22]. Participants were asked to find a quiet space for the session, use a computer or tablet with a minimum screen size of 10 inches, and wear headphones to decrease the impacts of environmental noises. During the session, the facilitator played the scenarios for the participant via screen-sharing. Between each scenario, participants switched to their own screen or a second device to access the questionnaire online.

For both conditions, the facilitator briefed participants about the study before proceeding with data collection. The presentation order of scenarios was randomised to overcome issues of fatigue and priming. The scenes in every scenario played automatically following a timer to ensure that each participant was exposed to the stimulus for the same amount of time. Each session, whether online or on-site, took an average of 35 minutes to complete.

Quantitative analyses (ANOVA, regression, and correlation) were conducted using IBM SPSS Statistics, while qualitative analysis (thematic analysis) was conducted using Microsoft Excel with two researchers for inter-rater reliability.

3.4 Participant Sample

Recruitment targeted 18–40-year-old subjects (mean = 26.34, SD = 5.19) who had lived in Singapore for at least two years for a more homogenous sample (due to the relatively small target sample size) which was familiar with the local public transport context. A total of 62 valid responses (female = 37) were accepted for analysis. Thirty-two participants took part in the on-site condition and 30 participants took part in the online condition. Participants received a cash voucher as a token of appreciation for their time. The study protocol was approved under NTU-IRB-2019-08-017 in Singapore.
4. RESULTS

4.1 Variables Influencing Companion Ratings

4.1.1 Emotional Responses

Figure 3 summarises participants’ emotional responses in a scatterplot. Using a scale of -5 to +5, the mean baseline pleasure and arousal levels were 1.47 and 0.77 respectively. After interacting with the Robot Companion in the commuter (CS), emergency (ES) and tourist/unfamiliar (TUS) scenarios, the average pleasure levels of participants were 0.63, 1.18, and 1.55 respectively, while the average arousal levels were 5.53, 6.10, and 6.37 respectively. Interaction with the Virtual Companion yielded higher pleasure levels of 2.65, 2.05, and 2.40 and higher arousal levels of 1.74, 1.81, and 1.45 respectively.

A 2-way within-subjects analysis of variance (ANOVA) was used to investigate if there were any differences in pleasure ratings (Figure 4) between Companion types (Robot Companion and Virtual Companion) across the three scenarios (CS, ES and TUS). The main effect of Companion type was significant, $F(1, 61) = 30.93, p < .001, \eta^2 = 0.34$, where the Virtual Companion was found to elicit significantly higher levels of pleasure ($M = 2.37, SD = 0.2$) than the Robot Companion ($M = 1.12, SD = 0.21$). Meanwhile, the main effect of scenario was not statistically significant, $F(2, 122) = 2.26, p = .11$. The interaction effect of Companion types and scenarios was significant, $F(2, 122) = 4.2, p < .05, \eta^2 = 0.06$, where the difference in pleasure elicitation was greater in the commuter scenario than in the other two scenarios.

For arousal ratings (Figure 5), a 2-way within-subjects ANOVA was again used to investigate if there were any differences between Companion types across the three scenarios. The main effect of Companion type was significant, $F(1, 61) = 5.33, p < .05, \eta^2 = 0.08$, where the Virtual Companion was found to elicit significantly higher levels of arousal ($M = 1.67, SD = 0.18$) than the Robot Companion ($M = 1, SD = 0.24$). Meanwhile, the main effect of scenario was not statistically significant, $F(2, 122) = 1.48, p = .23$. The interaction effect of Companion types and scenarios was significant, $F(2, 122) = 3.79, p < .05, \eta^2 = 0.06$, where the Virtual Companion elicited greater levels of arousal than the Robot Companion only in the commuter and emergency scenarios.

Based on the positive emotions generated from interaction with the Virtual Companion, H1a (the presence of must-haves will not cause positive emotion) is rejected and H1b (the absence of ideal-to-haves will not cause negative emotion) is supported. H2 (the absence of must-haves will cause negative emotion) appears to be supported and H3 (the presence of ideal-to-haves will cause positive emotions) rejected, with mean pleasure ratings for the Robot Companion falling below baseline.

H4 (scenarios influence subjective evaluations of function) is partially supported, applying only to the commuter and emergency scenarios and not to the tourist/unfamiliar scenario.
Table 1: Correlations between participants’ comfort with technology and trait ratings

| Comfort with Technology | r(60) |
|-------------------------|-------|
| Tourist / Unfamiliar Scenario | | |
| Trustworthiness | 0.27* |
| Comfort | 0.27* |
| Satisfaction | 0.27* |
| Virtual Companion | | |
| Appraochableness | 0.27* |
| Emergency Scenario | | |
| Arousal | 0.26* |
| Trustworthiness | 0.29* |
| Comfort | 0.26* |
| Tourist / Unfamiliar Scenario | | |
| Pleasure | 0.29* |
| Arousal | 0.25* |
| Trustworthiness | 0.25* |
| Comfort | 0.28* |

Note: * indicates $p<.05$, ** indicates $p<.01$

4.1.2 Comfort with Technology and Trait Ratings

To assess the linear relationship between comfort with technology and perceptions of the Virtual and Robot Companions, a bivariate Pearson’s product-moment correlation coefficient ($r$) was calculated. Statistically significant results are shown in Table 1. Results suggest that greater comfort with technology correlates positively with more positive evaluations of the Companions, supporting H5a.

4.1.3 Subject Personality, Ratings and Preferences

A Pearson correlation test was conducted to examine the correlation between participant personality and perceptions of the Companions. Only extraversion consistently correlated with various measures across all three scenarios. Statistically significant results are shown in Table 2. Extraversion is positively correlated with perceived Virtual Companion approachableness across all three scenarios. This indicates that an introverted individual will likely find the Virtual Companion less approachable. On the other hand, neuroticism was found to be negatively correlated with several Virtual Companion trait ratings. Based on these results, H5c is rejected.

Additionally, a logistic regression was carried out to test whether personality traits can predict participants’ preferred Companion type. Further supporting the previous finding, neuroticism accounted for 9.8% of the variance of the preferred Companion in the tourist/unfamiliar scenario. The unstandardized Beta weight for the constant is $B = -3.49$, $SE = 1.25$, $Wald = 7.82$, $p < .01$. The unstandardized Beta weight for neuroticism is $B = .80$, $SE = 0.36$, $Wald = 4.84$, $p < .05$. Participants are approximately 123.0% [Exp($B=2.23$, 95% CI[1.35, 7.58])] more likely to prefer the Robot Companion in the tourist/unfamiliar scenario with each one unit increase of their neuroticism score.

4.1.4 Gender, Emotional Response and Trait Ratings

A one-way between groups ANOVA was used to investigate if there were any differences in Companion ratings between genders. In the commuter scenarios, the ANOVA was statistically significant in various measures, indicating that males found the Robot Companion more trustworthy ($F(1,60) = 8.03$, $p < .01$, $\eta^2 = 0.13$) and comforting ($F(1,60) = 5.04$, $p < .05$, $\eta^2 = 0.08$) than females. In the emergency scenario comparing the gender groups, the ANOVA was also statistically significant in various measures, indicating that males experienced greater arousal ($F(1,60) = 5.04$, $p < .05$, $\eta^2 = 0.08$), found the Robot Companion more trustworthy ($F(1,60) = 6.36$, $p < .05$, $\eta^2 = 0.11$) and comforting ($F(1,60) = 7.27$, $p < .01$, $\eta^2 = 0.12$), and are more satisfied than females ($F(1,60) = 4.32$, $p < .05$, $\eta^2 = 0.07$). In the tourist/unfamiliar scenario, ratings of the Companion did not differ between genders. We note that differences identified in the commuter scenario and emergency scenario were only applicable to the Robot Companion, but no differences in Virtual Companion ratings were observed across genders. These differences could be due to males’ greater comfort with technology.
indicating a possible link between H5a and H5b. Their comfort with technology scores were higher compared to females $F(1, 60) = 4.21, p < .05, \eta^2 = 0.07$.

Based on the results, H5b is supported for the Robot Companion but not for the Virtual Companion.

4.2 Understanding the Rationale Behind Companion Preferences

Answers to the open-ended question on participants’ rationales for their Companion choice fell into several themes. Companion function played a major role. As expected, the Virtual Companion was perceived as better able to fulfill their mobility needs, i.e., must-haves, in all scenarios by providing pertinent, timely information and support. Participants said that they could then plan alternative routes or use their waiting time more meaningfully. This preference was especially critical for the commuter scenario, where subjects shared that their primary goal was to reach their destination as soon as possible. The Virtual Companion was also perceived as better able to serve more people at once than the Robot Companion. In the emergency scenario, participants overwhelmingly preferred the Virtual Companion. Its ability to make emergency calls and locate the vehicle were highly valued as they perceived quick response times as critical. A handful of subjects also considered a human touch important, referring to the human service personnel who appeared via the Virtual Companion.

Other reasons for preferring the Virtual Companion were based more on perceived weaknesses of the Robot Companion. For some subjects, the proactive greeting as they approached the bus stop or vehicle was seen as intrusive, “weird”, and even “creepy”, especially since the Robot Companion recognized them and their travel habits. Several stated that they preferred if the Robot Companion did not approach them first, i.e., the subjects wanted to initiate interactions. A few participants found the use of speech as the communication mode a privacy concern, describing how passengers in their vicinity would be able to overhear the conversation.

The physical assistance provided by the Robot Companion was not seen as particularly useful by most participants in the commuter scenario. Some stated that they could leave their heavy belongings on the ground instead. On the other hand, the physical guidance provided by the Robot Companion in the tourist/unfamiliar scenario was perceived as an advantage by almost a fifth of participants. Several subjects mentioned how the presence of the Robot Companion was comforting and gave a more personal touch, while others felt that it was reassuring in a crowded space, particularly among subjects high in neuroticism. They said: “It can physically accompany and direct me to the location of the bus. This is reassuring especially when it is packed with people” and “Robot Companion will be able to provide me with immediate assistance and comfort that I need especially given how crowded the interchange is and I probably do not have the patience to use the Virtual Companion.” A few participants thought that navigation was easier and less confusing with the Robot Companion guiding them, with one subject opining that this was especially useful for older passengers who may not be as tech-savvy with using the Virtual Companion. A similar number of participants found the Robot Companion useful for the emergency scenario, appreciating the function of physical support for the unwell, elderly passenger. However, while a few participants acknowledged this function, they still preferred the Virtual Companion due to its ability to call for emergency help.

4.3 Methodological Differences

Multiple one-way between groups ANOVAs were also used to investigate if ratings of the Companions differed based on the data collection condition. Interestingly, ratings of the Robot Companion differed on several measures in the commuter and emergency scenarios. In the commuter scenario, the Robot Companion was rated more useful and trustworthy by participants who completed the study on-site, versus participants who did the study online ($F(1, 60) = 4.40, p < .05, \eta^2 = 0.07$; $F(1, 60) = 4.79, p < .05, \eta^2 = 0.08$). Meanwhile, in the emergency scenario, the Robot Companion was rated more trustworthy by on-site participants ($F(1, 60) = 9.56, p < .01, \eta^2 = 0.16$) and elicited greater arousal than in online participants ($F(1, 60) = 5.15, p < .05, \eta^2 = 0.09$).

5. DISCUSSION, LIMITATIONS AND FURTHER RESEARCH

5.1 Deviations from the Kano Model

The results suggest that the link between emotional responses, must-haves and ideal-to-haves does not align completely with the Kano model. This could be because the present study adds emotion as a dependent variable, while studies based on the Kano model usually measure subjective evaluations through satisfaction. Studies have shown that the way key Kano relationships are expressed linguistically influences results [23]. Next, the method of depicting each Companion’s functions through different embodiments and the method of rating based on emotional...
scales could have confounded results. The Kano model method usually poses two questions, the functional and dysfunctional, to subjects after exposing them to specific product attributes [16]. Our study presented the two concepts in their entirety rather than breaking them down into specific characteristics, which could have led to the misattribution of the sources of positive or negative emotion [24]. Further isolation of variables could improve the precision of results.

Viewed through the lens of task completion [18] and need fulfilment [19] however, shows that our findings fall in line with expected performance. The familiarity and consequent usability of the Virtual Companion as a text-based touchscreen coupled with the fulfilment of participants’ primary goal, e.g., to arrive quickly at their destination, led to positive pleasure responses. On the other hand, the apparent lack of function in the Robot Companion, as evidenced by feelings of frustration or disappointment shared in the qualitative responses, and perceived invasion of privacy probably led to lower pleasure ratings.

5.2 Subject Backgrounds Affect Preferences

Participants who were more comfortable with new technology tended to rate the Robot and Virtual Companions more favourably on various measures of emotion (pleasure and arousal), usefulness, trustworthiness, comfort, and satisfaction. This could be due to a higher level of confidence towards technology, influencing one’s receptivity towards it.

Personality correlated to the perceptions of two Companion’s traits. Extraversion was linked with more favourable ratings for both the Virtual and Robot Companions. Research has found strong associations between the extraversion personality trait and sensation-seeking [25], defined as “the need for varied, novel, and complex sensations and experiences and the willingness to take physical and social risks for the sake of such experience” [26, p.10]. Meanwhile, participants scoring higher on neuroticism preferred the Robot Companion more than participants who were less neurotic, but only in the tourist/unfamiliar scenario. The scenario might have elicited higher levels of anxiety and stress, causing them to be more inclined towards the Robot Companion which possesses a more human-like touch.

Finally, gender played a role as males were found to be more comfortable and confident with the technology of the Virtual and Robot Companions. This finding is supported by existing literature where males have been found to be more confident in adopting technology for education purposes [27]. In a similar study, where gender differences were identified in career paths (i.e., males are more inclined to technology and science and females to humanities and social sciences), Trusz [28] cited the gender stereotype threat and self-fulfilling prophecy as two explanations for the phenomenon: Gender-based expectations from others self-propelled individuals towards the respective course choices. In our context, males’ confidence with technology could have influenced greater receptivity towards new technology, such as the Companion HMI, and eventually positively influence their ratings.

5.3 Imitating Human-Human Interactions – Or Not

Bus drivers today fulfil the roles of greeting passengers, where they might recognize regular commuters, and answering questions about bus services and routes. In our study, however, proactive greetings and personalized conversations with the Robot Companion were generally poorly received, which contrasts with the findings of Zhang et al. [29] showing positive valence and arousal with the addition of human-like features, and Hamacher et al. [30], who showed that people preferred a personable, communicative assistive robot over an efficient and reliable one. More research into the effect of demographic background, sociocultural contexts [31] and the Uncanny Valley [32] could be conducted to understand this difference. Further considerations are how anthropomorphic features in robot appearances have been found to raise expectations of their functional capabilities [33] and prompt people to base their interactions with robots off human-human social scripts, leading to further expectancy bias [34]. The nature of sharing information with the Robot Companion, essentially a digital interface, might also be perceived through the lens of data privacy. Subjects did not appear to question sharing route information with the Virtual Companion, perhaps due to familiarity with touchscreen directories. Another consideration is the setting of the HMI. Our study is based on shared spaces onboard public autonomous buses and stops. Research and development into automotive user interfaces (UI) tends to focus on private vehicles as in [35, 36], and these trend toward personalised, concierge-based services, such as Toyota’s YUI [37].

5.4 Methodological Reflections

Ratings for usefulness, trustworthiness, and emotional responses towards the Robot Companion in the commuter and emergency scenarios were higher from on-site participants compared to online participants. Similar find-
ings can be observed in the literature, as in a study conducted by Ward et al. [38], where the anonymity of online questionnaires allowed participants to be more honest in their responses, compared to pen-and-paper surveys. Priming might also have inadvertently played a part with the two data collection conditions. The on-site study in lab conditions created a more sterile environment which was detached from the familiar surroundings of participants' homes, leading to the two HMI concepts framed and thus received differently. This could have led to a subconscious increase in participants' trust towards the more novel Robot Companion. Another limitation concerns self-selection bias, which could have resulted in subjects with more interest or comfort with technology signing up for the study. Finally, participant recruitment focused on individuals between the ages 18–40. Further work could gather data from other age groups to extend generalisability.

6. CONCLUSION

In this study, two HMI concepts for autonomous buses were evaluated to determine how must-have and ideal-to-have functions in various scenarios affect user emotion and perception. The findings have implications for the design of HMI for future automated services, particularly conversational user interfaces, not just in mobility but also in other application domains. It would be prudent to design conversational interfaces in shared spaces carefully and not simply model them after human-human interactions. Additionally, identifying specific users is crucial in deciding which interaction mode to use, since findings show subject personality, gender, and comfort with technology played a role in how they rated the Companion. As one subject suggested, less tech-savvy people might find it easier or preferable to use conversational HMI which is tailored to their individual needs and habits. Methodologically, the site of data collection, i.e., on-site vs. remote, appeared to influence results. Future work should consider deeper investigation into specific design characteristics to determine the relationship between HMI concepts and emotional response more precisely.

ACKNOWLEDGEMENTS

This study is part of a research programme, TUMCREATE Phase 2, funded by the National Research Foundation (NRF) Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

REFERENCES

1. Kaasinen, E., et al.; Defining user experience goals to guide the design of industrial systems, Behaviour & Information Technology, 34(10), pp.976-991, 2015, doi: 10.1080/0144929X.2015.1035335.
2. Buchanan, R.; Wicked Problems in design thinking, Design Issues, 8(2), p.5, 1992, doi: 10.2307/1511637.
3. Land Transport Authority; Autonomous vehicles, LTA Industry & Innovations, 2019. https://www.lta.gov.sg/content/ltagov/en/industry_innovations/technologies/autonomous_vehicles.html (accessed 2021.03.01).
4. Guell, C., Panter, J., Jones, N.R., and Ogilvie, D.; Towards a differentiated understanding of active travel behaviour: Using social theory to explore everyday commuting, Social Science and Medicine, 75(1), pp.233-239, 2012, doi:10.1016/j.socscimed.2012.01.038.
5. Bissell, D.; Passenger mobilities: Affective atmospheres and the sociality of public transport, Environment and Planning D: Society and Space, 28(2), pp.270-289, 2010, doi: 10.1068/d3909.
6. Kurniawan, J. H., Ong, C., and Cheah, L.; Examining values and influences affecting public expectations of future urban mobility: A Singapore case study, Transport Policy, 66, pp.66-75, 2018.
7. Kong, P., Cornet, H., and Frenkler, F.; Designing for delight and disappointment with two companions for future mobility in Singapore, Proceedings of the 8th International Conference on Kansei Engineering and Emotion Research, 2020.
8. Russell, J. A.; Core affect and the psychological construction of emotion, Psychological Review, 110(1), pp.145-172, 2003, doi: 10.1037/0033-295X.110.1.145.
9. Kouprie, M., and Sleeswijk Visser, F.; A framework for empathy in design: Stepping into and out of the user’s life, Journal of Engineering Design, 20(5), pp.437-448, 2009, doi:10.1080/09544820902875033.
10. Dunn, J.R., and Schweitzer, M.E.; Feeling and believing: The influence of emotion on trust, 88(5), pp.736-748, 2005, doi: 10.1037/0022-3514.88.5.736.
11. Lottridge, D., Chignell, M., and Jovicic, A.; Affective interaction: Understanding, evaluating, and designing for human emotion, Reviews of Human Factors and Ergonomics, 7(1), pp.197-217, 2011.
12. Van Hagen, M., and Bron, P.; Enhancing the experience of the train journey: Changing the focus from satisfaction to emotional experience of customers, Transportation Research Procedia, 1(1), pp.253-263, 2014, doi: 10.1016/j.trpro.2014.07.025.
13. Kahneman, D.; Evaluation by moments: Past and
future, In Kahneman, D., and Tversky, A. (eds) Choices, values, and frames, Cambridge University Press, pp.693-708, 2000.

14. Carreira, R., Patrício, L., Jorge, R.N., and Magee, C.L.; Development of an extended Kansei engineering method to incorporate experience requirements in product-service system design, Journal of Engineering Design, 24(10), pp.738-764, 2013, doi: 10.1080/09544828.2013.834038.

15. Kong, P., Cornet, H., and Frenkler, F.; Closing the human-machine disconnect: Design requirements for two extreme companions for future autonomous mobility, Proceedings of the DESIGN 2020 16th International Design Conference, pp.1627-1636, doi: 10.1017/dsd.2020.92.

16. Kano, N., Seraku, N., Takahashi, F., and Tsuji, S.; Attractive quality and must-be quality, Journal of The Japanese Society for Quality Control, 14(2), pp.147-156, 1984, doi: 10.20684/quality.14.2.147. (in Japanese)

17. Syrdal, D. S., Dautenhahn, K., Woods, S. N., Walters, M. L., and Koay, K. L.; Looking good? Appearance preferences and robot personality inferences at zero acquaintance. AAAI Spring Symposium - Technical Report, SS-07-07, pp.86-92, 2007, [Online]. Available: https://www.aaai.org/Papers/Symposia/Spring/2007/SS-07-07/SS-07-07-019.pdf.

18. Jokinen, J. P.P.; Emotional user experience: Traits, events, and states, International Journal of Human Computer Studies, 76, pp.67-77, 2015, doi: 10.1016/j.ijhcs.2014.12.006.

19. Hassenzahl, M., Wiklund-Engblom, A., Bengs, A., Hägglund, S., and Diefenbach, S.; Experience-oriented and product-oriented evaluation: Psychological need fulfillment, positive affect, and product perception, International Journal of Human-Computer Interaction, 31(8), pp.530-544, 2015, doi: 10.1080/10447318.2015.1064664.

20. Rammstedt, B., and John, O. P.; Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German, Journal of Research in Personality, 41(1), pp.203-212, 2007, doi: 10.1016/j.jrp.2006.02.001.

21. Agarwal, R., and Prasad, J.; A conceptual and operational definition of personal innovativeness in the domain of information technology, Information Systems Research, 9(2), pp.204-215, 1998, doi: 10.1287/isre.9.2.204.

22. Zoom Video Communications Inc.; Zoom, 2020. https://zoom.us/ (accessed 2020.12.08).

23. Horton, G., and Goers, J.; A revised Kano model and its application in product feature discovery, University of Magdeburg, Germany, 2019. [Online]. Available: https://www.researchgate.net/publication/332304132_A_Revised_Kano_Model_and_its_Application_in_Product_Feature_Discovery (accessed 2019.10.21).

24. Alba, J. W., and Williams, E. F.; Pleasure principles: A review of research on hedonic consumption, Journal of Consumer Psychology, 23(1), pp.2-18, 2013, doi: 10.1016/j.jcps.2012.07.003.

25. Aluja, A., García, Ó., and García, L. F.; Relationships among extraversion, openness to experience, and sensation seeking, Personality and Individual Differences, 35(3), pp.671-680, 2003, doi: https://doi.org/10.1016/S0191-8869(02)00244-1.

26. Zuckerman, M.; Sensation seeking: Beyond the optimal level of arousal, Hillsdale, N.J.: L. Erlbaum Associates, New York, 1979.

27. You, H. K., and Cheng, A. L. F.; Gender difference of confidence in using technology for learning, The Journal of Technology Studies, 38(2), pp.74-79, 2012, doi: 10.21061/jots.v38i2.a.2.

28. Trusz, S.; Why do females choose to study humanities or social sciences, while males prefer technology or science? Some intrapersonal and interpersonal predictors, Social Psychology of Education, 23(3), pp.615-639, 2020, doi: 10.1007/s11218-020-09551-5.

29. Zhang, T., Kaber, D. B., Zhu, B., Swangnetr, M., Mosaly, P., and Hodge, L.; Service robot feature design effects on user perceptions and emotional responses, Intelligent Service Robots, 3(2), pp.73-88, 2010, doi: 10.1007/s11370-010-0060-9.

30. Hamacher, A., Bianchi-Berthouze, N., Pipe, A. G., and Eder, K.; Believing in BERT: Using expressive communication to enhance trust and counteract operational error in physical Human-Robot interaction, Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp.493-500, 2016, doi: 10.1109/ROMAN.2016.7745163.

31. Złotowski, J., Proudfoot, D., Yogeeswaran, K., and Bartneck, C.; Anthropomorphism: Opportunities and challenges in Human–Robot interaction, International Journal of Social Robotics, 7(3), pp.347-360, 2015, doi: 10.1007/s12369-014-0267-6.

32. Mori, M.; The uncanny valley, Energy, 7(4), pp.33-35, 1970, [Online]. Available: http://www.movingimages.info/digitalmedia/wp-content/uploads/2010/06/MorUnc.pdf (Accessed 2018.05.10).
User Emotion and Perception of Virtual and Robot Companion Concepts for Shared Autonomous Mobility in Singapore

33. Haring, K. S., Silvera-Tawil, D., Takahashi, T., Watanabe, K., and Velonaki, M.; How people perceive different robot types: A direct comparison of an android, humanoid, and non-biomimetic robot, Proceedings of the 8th International Conference on Knowledge and Smart Technology (KST), pp.265-270, 2016, doi: 10.1109/KST.2016.7440504.

34. Edwards, A., Edwards, C., Westerman, D., and Spence, P. R.; Initial expectations, interactions, and beyond with social robots, Computers in Human Behavior, 90, pp.308-314, 2019, doi: 10.1016/j.chb.2018.08.042.

35. Pettersson, I., and Ju, W.; Design techniques for exploring automotive interaction in the drive towards automation, Proceedings of the 2017 Conference on Designing Interactive Systems - DIS '17, pp.147-160, 2017, doi: 10.1145/3064663.3064666.

36. Braun, M., Pfleging, B., and Alt, F.; At your service: Designing voice assistant personalities to improve automotive user interfaces a real world driving study, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp.1-11, 2019.

37. Etherington, D.; Here’s what it’s like to drive with Toyota’s Yui AI in-car assistant, TechCrunch, 2017. https://techcrunch.com/2017/01/06/heres-what-its-like-to-drive-with-toyotas-yui-ai-in-car-assistant/ (accessed 2018.04.19).

38. Ward, P., Clark, T., Zabriskie, R., and Morris, T.; Paper/pencil versus online data collection: An exploratory study, Journal of Leisure Research, 44(4), pp.507-530, 2012, doi: 10.1080/00222216.2012.11950276.

Penny KONG (Non-member)
Penny Kong was a Research Associate in the Design for Autonomous Mobility (DAM) team at TUMCREATE. Currently, she leads design and research activities at TUM Asia for the project “2025 and Beyond: Re-defining People-Vehicle Interaction and Infrastructure for Future Mobility in Singapore” under the LTA-URA Urban Mobility Grand Challenge Grant Call for Designing Bus Stops, Taxi Stands and Pick-Up Drop-Offs for an Autonomous Vehicle Future. Her doctoral research investigates the design of human-machine interfaces in shared autonomous vehicles for affective trust.

Pei Yi LIM (Non-member)
Lim Pei Yi was a Project Officer in the Design for Autonomous Mobility (DAM) team at TUMCREATE. She is currently a Research Assistant at Thye Hua Kwan Moral Charities (THKMC) where she rigorously trials, tests, and evaluates various programmes for children with developmental needs and ensures that they are delivering the best possible outcomes for them. Research findings are directly translated into real-life practice, with an aim to ensure programme fidelity and continuous quality improvement.

Henriette CORNET (Non-member)
Dr. Henriette Cornet was Principal Investigator of the Design for Autonomous Mobility (DAM) team at TUMCREATE. She is currently Senior Manager – Knowledge & Innovation at UITP (International Association of Public Transport), where she coordinates activities related to Connected Cooperative and Automated Mobility under the EU-funded project “SHOW SHared automation Operating models for Worldwide adoption” aiming at deploying fleets of shared AVs in 20 cities in Europe with 70 consortium partners.

Fritz FRENKLER (Non-member)
Fritz Frenkler was Full Professor of Industrial Design at the Technical University of Munich (TUM) from 2006 until 2020 and Programme Principal Investigator of the Design for Autonomous Mobility (DAM) team at TUMCREATE. Since October 2020, he is TUM Emeritus of Excellence at the TUM Senior Excellence Faculty. Prof. Frenkler is CEO and co-founder of f/p design, a design company with offices in Berlin, Munich and Kyoto. Starting from 1996, he is Member of the Board for iF Industrie Forum Design e.V. (since April 2018, iF Design Foundation) and from 2013, he is member of the Akademie der Kuenste (AdK), Berlin.