Conducting HRI Research 'In the Wild': Reflections Based on Two Urban Robot Studies

Marius Hoggenmueller
Urban Interfaces lab, Sydney School of Architecture, Design and Planning
The University of Sydney
Sydney, Australia
marius.hoggenmueller@sydney.edu.au

***Abstract*—In this paper, I present two case studies on urban robots that were both conducted 'in the wild' but differed in terms of the role of design, research methods, and deployment context. I provide critical reflections on the different approaches undertaken in those two studies, and how they impacted scope and results. Furthermore, I argue that researching HRI in the wild requires qualitative in addition to quantitative approaches, and why we should study robotic artefacts in the wild early on.**

***Index Terms*—human-robot interaction, research in the wild, urban robots, design research, research methods**

I. Introduction

There have been various discussions in the community on how we research human-robot interaction (HRI). For example, some researchers have recently argued for designerly HRI work as a legitimate form of knowledge production [1]; some contextualised research through design approaches, which are already widely adopted by human-computer interaction (HCI) researchers [2], in HRI [3]; and others reviewed the usage of qualitative research approaches in HRI [4]. Also, there have been workshop series that focus on the study of HRI in the 'real-world', such as public spaces [5]. While varying in terms of the proposed approaches, most of those efforts are motivated by the observation that the currently predominant type of hypothesis-driven study, often conducted in the lab and using quantitative measures, has limitations: for example, when it comes to designing robots for a specific context, capturing the experiential qualities of interaction in the 'real-world' [6], or to explore what is the right thing to design [3].

In 2019, our research team (myself together with researchers from the Urban Interfaces lab) conducted two in the wild studies [7] on urban robots in the greater area of Sydney, Australia. While the two studies involved the same robotic artefact called Woodie, they differed considerably in terms of motivation and assumptions, investigated research questions, methodological considerations, and the socio-spatial context of the deployment locations. In this paper, I do not intend to provide detailed account to the specific results of each study (which can be found in the original publications [8], [9]); rather, I provide critical reflections on the undertaken approaches within those in the wild studies, and how they impacted scope and results. Furthermore, I argue that researching HRI in the wild requires 'balanced' mixed-method studies and why we should study robotic artefacts in the wild early on.

Fig. 1. First in the wild study with the urban robot Woodie as part of an annual lighting festival. More pictures from the deployment can be found in the original publication [8] and on https://marius.hoggenmueller.com/woodie/.

II. Two in the Wild Studies

In the following, I briefly introduce the two in the wild studies that informed the subsequently following reflections. The two studies were conducted within the larger scope of my PhD research, which aims to advance the understanding of the design space of urban robotic interfaces. This overview is far from exhaustive, but should provide the reader with sufficient background information about the urban robotic artefact and the conducted studies.

A. Urban Robots to Enable Physicalised Displays

The first study set out to investigate a novel form of pervasive urban display [10], which produces content in a physicalised form. Building on previous research highlighting the experiential and transient qualities of non-digital displays (e.g. public visualisations using chalk [11]), the aim was to replicate and automate the same through a self-moving robotic platform. In a research through design approach [2] that was oriented towards more generative and speculative interpretations (c.f. [12], [13]), we created Woodie: a slow-moving urban robot capable of drawing with chalk sticks on the ground [8]. In order to communicate to surrounding pedestrians, a low-resolution (low-res) lighting display was integrated into the robot's outer shell. Woodie was designed for an annual lighting festival based in Sydney, featuring installations across the whole city. Over three weeks, the robot was deployed in

a quiet laneway in a densely populated northern suburb of Sydney (see Figure 1). In an area of roughly 15 metres in length and 5 metres in width, Woodie was drawing pre-programmed visualisations, such as flowers and love hearts. The area was illuminated by ultraviolet (UV) lamps, and the robot was drawing with luminescent chalk sticks, which increased visibility of the drawings and created a sublime atmosphere. On most evenings, chalk sticks were handed out to visitors, so they could extend Woodie's drawings or add their own. During the deployment, we collected data through interviews (N=21), observation notes, and video recordings.

Our data analysis revealed that Woodie successfully attracted passers-by and acted as a facilitator to engage people in collaborative and creative placemaking. This was manifested through people standing in a circle around Woodie as it were a street performer, wandering around the place to read other people's drawings, adding their own drawings, and children would often lie on the floor to see the chalk stick touching the ground to find out more about the drawing mechanism. Our analysis revealed various audience behaviour patterns around the installation, which were influenced by the number of people on-site and whether only the robot was drawing (resulting in focused audience grouping around the robot) or whether chalk sticks were handed out to people and therefore most of the laneway was covered with drawings (resulting in diffused audience grouping around the whole area). Despite Woodie not being a humanoid robot and the absence of any emotional expressions or feedback, the findings suggested that people had an emotional response towards the robot and perceived it as a living being. For example, Woodie was often referred to as "something cute", akin to "living organisms" such as a "mushroom" or a "jellyfish". People would also speculate about gender and character traits, e.g. asking festival staff if Woodie was "male or female", or "serious".

B. Emotional Expressions of Non-Humanoid Urban Robots

In the subsequent study, we investigated, in a more focused manner, the ability of non-humanoid urban robots to express artificial emotions. In HRI, there has been a line of research that investigated various abstract information cues on robots (e.g. motion, colour, sound) for affective communication (see meta review in [14]). The research is based on the assumption that affective behaviour can increase social acceptance of robots and that humans tend to anthropomorphise robots [15]. Several studies have shown that people can correctly interpret artificial emotional expressions to a certain extent. However, all of the previous studies were conducted in controlled lab environments. To explore if the expression of artificial emotions would stand the 'messiness' of an urban situation, we designed and implemented six basic emotions for Woodie (expressed through light patterns and movement) and deployed the robot in a pedestrianised area of our university's campus [9] (see Figure 2). Other than in the initial study, we diminished the robot's drawing feature and did not alter the surrounding environment in order to solely focus on the impact of the robot's emotional expressions.

Fig. 2. Second deployment of Woodie in a pedestrianised area to investigate emotional expressions of non-humanoid urban robots.

Following a mixed-method evaluation approach, we asked passers-by (N=72) about their first impression, then handed out a questionnaire for them to rate to what extent they agreed that the displayed emotion corresponded to each of the six implemented emotions, followed by a short interview to better understand the reasons for their ratings.

Quantitative results from this study suggested that participants had difficulties interpreting the displayed emotions correctly. Even for displayed emotions that were ranked highest in their corresponding category (e.g. anger, sadness), these emotions would still be ranked high in some of the other categories as well. Interestingly, the pre-questionnaire interviews revealed that only 2 out of the 72 interviewees initially related the robot's behaviour to emotions. Based on the analysis of post-questionnaire interviews, we found that the form factor of the robot (e.g. its shape) influenced perception and interpretation. Furthermore, we identified various influencing contextual aspects that are extrinsic to the urban robot itself. We found that the overall choreography and interactional context had an influence on people's interpretations. For example, people attributed low-valence emotions (e.g. sad, nervous) to the robot based on its repetitive movement (i.e. performing the same emotional expression in a loop) and its movement radius (which was considered rather limited in comparison to the spacious layout of the deployment location). Independent of the expressed emotion, also the immediate impact of the environment as well as people's own emotions and past experiences influenced perception and interpretation. This includes, for example, the ambient lighting, weather, and whether people associate the surrounding campus environment with positive or negative emotions.

III. Reflections

As previously stated, both of the here presented studies were part of my PhD research on urban robotic interfaces. Prior to my PhD, I conducted research on other urban technologies, mainly contributing to the field of media architecture [16]. Based on my background and lessons learnt from previous research projects in public space, I was convinced that I would

need to consider the situational context and impact of real-world environments when designing prototypes and evaluating them with users. Building on thoughts from HRI researchers (e.g. [5], [6]) that lab or online studies fall short of predicting "how real people, in real-world environments, would interact [...] with a real robot" [6], I further planned early on to conduct most of my studies in the wild [7]; also, in order to capture the experience-oriented aspects that are apparent in third paradigm HCI research [17], which I believed I would stringently follow in terms of perspectives and approaches. However, it is now – after synthesising findings across cases and comparing individual approaches in-depth and in retrospect – that I conclude that this was not always the case. In the following, I reflect on the different approaches within the two in the wild studies; further, I discuss the importance of 'balanced' mixed-methods studies in hypothesis-driven HRI research in the wild, and why we should study HRI in the wild earlier.

A. In the wild approaches vary, and so do their scope and results

For the first in the wild study on Woodie, we followed a research through design [2] approach. Similar to how Gaver et al. reflected on their design approach [12], our design was less informed by theory than concrete examples from design practice. This also included design work that might appear to have little to do with the HRI field: for example, the design of Woodie's outer shell was informed by the aesthetics of low-res media façades. However, not only for the design of the robot, but also how we contextually-embedded the robot within the surroundings, we drew on media architecture principles [16], for example, by enabling participatory experiences through handing out chalk sticks to people and by setting the architectural scene through the ambient UV lighting. Lastly, our evaluation focus on audience behaviour and engagement was motivated and methodologically inspired by previous research on public displays and media architecture. These considerations, the speculative tendencies and the deployment at a festival context, would probably qualify this project as what Zimmerman et al. [2] refer to as a 'showroom' approach. On the other hand, for the second in the wild study, our research approach was adopted from experimental psychology. Design itself played a secondary role; rather, design was a means to test our hypotheses. As such, our experimental design and procedure largely followed those of laboratory studies. Rather than setting the scene for a holistic narrative, we simply remote-controlled the emotional expressions and looped them until we reached enough participants per condition. The robot was simply deployed at the study location without further consideration of the socio-spatial context.

Although both studies were conducted in the wild, those differences in research approach and the varying roles of design clearly impacted the scope and results of the studies, not only in regards to what we learned as designers and researchers, but also how they shaped the user. In the first in the wild study, we better achieved to envision, design, and study the robot as a socio-material configuration [18] in the urban environment. I suspect that the experience would have been largely different if we had not let people draw with the robot or neglected the surrounding environment. As one participant expressed: "[...] the combination of we can draw ourselves, but the robot draws as well, is a nice touch. It makes it more human, I suppose, than only a robot [drawing]." Another participant, a mother of two young children, was clearly immersed in the speculative world that we created when she stated that "to [her] it's fantastic" to imagine that her children could be "entertained for hours" by robots in future cities. In comparison, in the second in the wild study, participants were more sceptical towards the robot. For example, they would speculate about its "functionality" and "purpose" as the robot did not interact with the surrounding environment or people. Some participants – after being briefed about the purpose of the study – would also question "[why] it [would be] necessary for a robot to express negative emotions in public spaces".

Certainly, the discussion of different research paradigms and approaches is not new and there is also existing literature from HCI as guidance that is applicable to the design of robots alike (e.g. [2]). However, I believe that the provided reflections from our two in the wild studies can still be helpful for other HRI researchers. In particular, studying robotics in the wild is rather new to the HRI field, and one could easily presume that by simply evaluating HRI in the 'real-world' (e.g. urban contexts), one could overcome the limitations of laboratory studies. This might be the case to a certain extent: indeed, we were able to identify contextual aspects that affect the interpretation of emotional expressions of urban robots, which probably would not have been possible to measure in the lab. However, comparing across both studies, it is my believe that for designing more desirable robots and envision them as "socio-material assemblages" [18], it is necessary to equally consider context, environment and people not only during the in-situ evaluation, but already in design and planning. To do so, it can also be beneficial to integrate conceptual and methodological knowledge from other fields, such as, in our case, media architecture.

B. A call for 'balanced' mixed-methods studies

Another reflection that I believe is worth sharing is about the importance of 'balanced' mixed-method approaches (i.e. quantitative and qualitative) in hypothesis-driven studies that are carried out in highly context-dependent settings (e.g. the urban context). In our second study on Woodie, we aimed to pay similar attention to quantitative and qualitative data. This means that we analysed the qualitative data with the same methods (e.g. thematic analysis) and rigour as if it would have been a purely qualitative study. While the quantitative results showed that people had difficulties in interpreting the emotions correctly, it was foremost the qualitative data which helped us identify reasons. It was then also the qualitative findings that let us formulate design considerations to support other designers and researchers on how to improve the design of emotional expressions for urban robots. While qualitative research studies are represented in the field, a recent review

by Veling and McGinn [4] has shown that only 7% and 5% of HRI research papers that used respectively interviews and observations are hypothesis-driven studies. This suggests that the collection of qualitative data is still rather underrepresented in hypothesis-driven studies on human-robot interaction. Further, their review stressed that reporting approaches and rigour in qualitative HRI studies differ substantially. However, based on the reflections from my own studies, I suggest that it is necessary, in particular for in the wild studies (given their highly context-dependent nature), to consider qualitative data as equally important as quantitative data.

C. Re-configuring design research practices

The last area of reflection concerns how and what we study in HRI, the way we currently do. Given that lab studies are often less expensive to run [7], we tend to conduct design research in the lab before eventually moving forward to in-situ evaluations. Certainly, this can be an appropriate approach, in particular for product development, as prototypes that are evaluated in the wild (and over longer periods of time) have higher requirements in terms of stability and functionality. However, I see the following issues in current HRI research: When a new research topic (such as the study of artificial emotions for non-humanoid robots, just to name one example) becomes prominent, we tend to quickly see a wide range of studies which investigate the same topic for different modes of communication, applied to different robot types, and so on. However, we see research in the wild or using contextualised prototyping approaches less frequently early on. To the best of my knowledge, our second study on Woodie was the first one of this kind to be conducted in a real urban environment. What I think is problematic here is that a considerable amount of time and resources are being spent before validating if the results are applicable to 'real-world' contexts. Furthermore, we often introduce and motivate our research (no matter if lab or in-situ) with platitudes, such as "robots are leaving factories" or "robots are entering our daily lives" (as recently brought up in a message on Twitter by Guy Hoffman[1]). In my opinion, this leads to two considerations: Firstly, we should consider testing our designs in the wild earlier; prototypes don't have to be 'complete', and approaches like urban probes [19] could be further exploited in HRI. Secondly, we need similar discussions as in HCI on the problem-solving capacity of our research [20]. For example, are our results transferable to those "robots entering our daily lives"? Or using artificial emotions as an example again: Are we studying a significant problem if humans might not want robots to express negative emotions in public space (as stated by some of our participants)?

[1] https://twitter.com/guyhoffman/status/1413903856367509518?s=20, *last accessed: January 2022*

References

[1] M. L. Lupetti, C. Zaga, and N. Cila, "Designerly ways of knowing in hri: Broadening the scope of design-oriented hri through the concept of intermediate-level knowledge," in *Proc. of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '21. New York, NY, USA: ACM, 2021, p. 389–398. [Online]. Available: https://doi.org/10.1145/3434073.3444668

[2] J. Zimmerman and J. Forlizzi, *Research Through Design in HCI*. New York, NY: Springer New York, 2014, pp. 167–189. [Online]. Available: https://doi.org/10.1007/978-1-4939-0378-8_8

[3] M. Luria, M. Hoggenmüller, W.-Y. Lee, L. Hespanhol, M. Jung, and J. Forlizzi, "Research through design approaches in human-robot interaction," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '21 Companion. New York, NY, USA: ACM, 2021, p. 685–687. [Online]. Available: https://doi.org/10.1145/3434074.3444868

[4] L. Veling and C. Mcginn, "Qualitative research in hri: A review and taxonomy," *International Journal of Social Robotics*, 02 2021.

[5] M. E. Foster, M. Giuliani, and R. Petrick, "Pubrob.org - the workshop & tutorial series on public space human-robot interaction," https://pubrob.org, *last accessed: January 2020*, 2011–2018.

[6] K. Dautenhahn, "Some brief thoughts on the past and future of human-robot interaction," *J. Hum.-Robot Interact.*, vol. 7, no. 1, may 2018. [Online]. Available: https://doi.org/10.1145/3209769

[7] A. Chamberlain, A. Crabtree, T. Rodden, M. Jones, and Y. Rogers, "Research in the wild: Understanding 'in the wild' approaches to design and development," in *Proc. of the Designing Interactive Systems Conference*, ser. DIS '12. New York, NY, USA: ACM, 2012, p. 795–796. [Online]. Available: https://dl.acm.org/doi/10.1145/2317956.2318078

[8] M. Hoggenmueller, L. Hespanhol, and M. Tomitsch, "Stop and smell the chalk flowers: A robotic probe for investigating urban interaction with physicalised displays," in *Proc. of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: ACM, 2020, p. 1–14. [Online]. Available: https://doi.org/10.1145/3313831.3376676

[9] M. Hoggenmueller, J. Chen, and L. Hespanhol, "Emotional expressions of non-humanoid urban robots: The role of contextual aspects on interpretations," in *Proc. of the 9TH ACM International Symposium on Pervasive Displays*, ser. PerDis '20. New York, NY, USA: ACM, 2020, p. 87–95. [Online]. Available: https://doi.org/10.1145/3393712.3395341

[10] M. Hoggenmueller, L. Hespanhol, A. Wiethoff, and M. Tomitsch, "Self-moving robots and pulverised urban displays: status quo, taxonomy, and challenges in emerging pervasive display research," *Personal and Ubiquitous Computing*, 09 2020.

[11] L. Koeman, V. Kalnikaitė, Y. Rogers, and J. Bird, "What chalk and tape can tell us: Lessons learnt for next generation urban displays," in *Proc. of The International Symposium on Pervasive Displays*, ser. PerDis '14. New York, NY, USA: ACM, 2014, pp. 130:130–130:135. [Online]. Available: http://doi.acm.org/10.1145/2611009.2611018

[12] B. Gaver and J. Bowers, "Annotated portfolios," *Interactions*, vol. 19, no. 4, p. 40–49, jul 2012. [Online]. Available: https://doi.org/10.1145/2212877.2212889

[13] J. Auger, "Living with robots: A speculative design approach," *J. Hum.-Robot Interact.*, vol. 3, no. 1, p. 20–42, Feb. 2014. [Online]. Available: https://doi.org/10.5898/JHRI.3.1.Auger

[14] Y. Wang, L. Hespanhol, and M. Tomitsch, "How can autonomous vehicles convey emotions to pedestrians? a review of emotionally expressive non-humanoid robots," *Multimodal Technologies and Interaction*, vol. 5, no. 12, 2021. [Online]. Available: https://www.mdpi.com/2414-4088/5/12/84

[15] C. Breazeal, "Emotion and sociable humanoid robots," *International Journal of Human-Computer Studies*, vol. 59, no. 1, pp. 119–155, 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581903000181

[16] A. Wiethoff and H. Hussmann, *Media architecture: using information and media as construction material*. De Gruyter, 2017, vol. 8.

[17] S. Harrison, D. Tatar, and P. Sengers, "The three paradigms of hci," in *Ext. Abstracts CHI 2007*, 01 2007.

[18] L. Suchman, *Human-Machine Reconfigurations: Plans and Situated Actions*, 2nd ed., ser. Learning in Doing: Social, Cognitive and Computational Perspectives. Cambridge University Press, 2006.

[19] E. Paulos and T. Jenkins, "Urban probes: Encountering our emerging urban atmospheres," in *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '05. New York, NY, USA: ACM, 2005, pp. 341–350. [Online]. Available: http://doi.acm.org/10.1145/1054972.1055020

[20] A. Oulasvirta and K. Hornbæk, "Hci research as problem-solving," in *Proc. of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, p. 4956–4967. [Online]. Available: https://doi.org/10.1145/2858036.2858283