Augmenting Gestural Interactions with Mid-Air Haptic Feedback: A Case Study of Mixed-Method Longitudinal UX-Testing in the Lab

Lawrence Van den Bogaert, Isa Rutten, and David Geerts

Abstract Ultrasound mid-air haptic feedback is a novel output technology that allows users to experience a sense of touch in mid-air on the unadorned palm and fingers of the hand. Even though a growing body of research has studied various aspects of the UX of mid-air haptics, little is known about what happens to the users' perception and experience after repeated use. The main reason for this is that today, mid-air haptic technology is not easily integrated in everyday devices (e.g. smartphones) nor widespread, making it difficult for it to be tested outside of a lab environment. This chapter describes the set-up of a longitudinal in-lab study, in which a mixed-method design was used to understand how the hedonic, pragmatic and emotional aspects of the UX of mid-air haptics changed over time. In eight sessions, spread over a five-week period, 31 participants interacted with a gesture-controlled home automation system augmented with mid-air haptic feedback. We report in this chapter on our participant recruitment and retention approach, the mixed-method set-up that was used, and (an excerpt of) the main results. Subsequently, we summarize best practices and propose suggestions for researchers who in the future intend to conduct a multimethod longitudinal study.

1 Introduction

As novel technologies emerge at a fast-paced rate and researchers' resources are generally limited, studying user experience or design aspects over a longer period of time is often a challenge in human–computer interaction (HCI). Especially when

L. Van den Bogaert (✉) · I. Rutten · D. Geerts
Meaningful Interactions Lab (mintlab), Parkstraat 45, 3000 Leuven, Belgium
e-mail: lawrence.vandenbogaert@kuleuven.be

I. Rutten
e-mail: isa.rutten@kuleuven.be

D. Geerts
e-mail: david.geerts@kuleuven.be

E. Karapanos et al. (eds.), *Advances in Longitudinal HCI Research*,
Human–Computer Interaction Series, https://doi.org/10.1007/978-3-030-67322-2_10

the technology or interface of interest cannot be taken home or incorporated in participants' everyday lives (e.g. installed as an app), longitudinal testing becomes cumbersome and impractical, both for the researcher as well as for the participant. Nevertheless, long-term user tests can indeed yield valuable or even unexpected insights, making them worthwhile.

This chapter reports on a case study that was part of a research project on mid-air haptics, where we undertook such a longitudinal user study. With the sense of touch becoming of increased interest in the HCI field, researchers have been seeking ways to convey touchless haptic feedback to users as an alternative for vibrotactile feedback. Ultrasound mid-air haptic feedback, often referred to as mid-air haptic feedback or simply mid-air haptics, does so by generating ultrasound pressure fields that actuate the sense of touch in the palm and fingers [1]. In addition to a relatively accurate localization error of less than 1 cm [2], mid-air haptics can provide multi-point feedback as well as different modulation frequencies, allowing for idiosyncratic touch sensations [3]. Particularly promising is the combination of this technology with gestural interfaces, which intrinsically lack any form of haptic feedback.

Today, mid-air haptic technology is not commonly integrated in everyday devices (e.g. smartphones) nor widespread, making it difficult for it to be tested outside of a lab environment. As a consequence, insights on the effects of its prolonged repeated use are scarce. Even though research shows that mid-air haptics augment the user experience when added to, e.g. movie experiences [4], VR [5] and car simulations [6], it is unclear whether these beneficial effects stand the test of time.

We conducted a longitudinal study with 31 participants who each engaged with mid-air haptics on 8 separate occasions over a 5-week period in order for us to gain a better understanding of the effects of repeated interaction with this novel technology. Because we only disposed of one mid-air haptic device—they are not commonly available and still expensive—and because this device can not simply be connected to other hardware, each session of the study was conducted in-lab, requiring a considerable engagement from participants. In this chapter, we focus on the methodological challenges we encountered during this study. We will share our findings in terms of recruitment, participant retention and overall study design. In summary, this chapter aims to report best practices and study set-up suggestions for HCI researchers who in the future intend to conduct in-lab longitudinal studies.

2 Related Work

2.1 Assessing the UX of Mid-Air Haptics

To evaluate a product's UX and design aspects, a variety of methods, assessment tools, frameworks and theories has been proposed and discussed in the HCI field. Hassenzahl's [7] framework on hedonic and pragmatic aspects of user experience is one of the most prevalent and distinguishes, on the one hand, between product

attributes that are connected to the user's need to achieve behavioural goals (pragmatic), such as performing a task effectively or understanding the functionalities of a product easily, and on the other hand, product attributes that are connected to the user's self (hedonic) such as the human need to express oneself through objects (identification). In line with Hassenzahl, Mahlke and Thüring's [8] holistic model of UX also identifies instrumental (or pragmatic) and non-instrumental (or hedonic) components of the UX, but adds a third component: the users' emotional reactions.

In their pursuit to 'measure the added value' of specifically mid-air haptics, Maggioni et al. [9] built on this three-part model. Maggioni et al. [9] too consider the dualistic pragmatic/hedonic nature of UX (mainly drawing on Hassenzahl's work and the related AttrakDiff questionnaire as an assessment tool for these two components) and add to that the assessment of the user's valence and arousal as indications of emotional reactions, in line with Mahlke and Thüring [8]. In addition, as a fourth component, they incorporate the potential effect of the user's pre-exposure expectations, as these have been shown to influence users' experiences [10].

Next to Maggioni et al. [9], other authors also gauged the effects and outcomes of adding mid-air haptics to an interface. Whereas Ablart et al. [4] found an increase in arousal and valence (i.e. emotional response) when mid-air haptics augmented a one-minute video viewing experience, Hwang et al. [5] showed an increase in enjoyment when playing a VR piano that was augmented with mid-air haptic feedback. Limerick et al. [11], in turn, used the User Engagement Scale (UES) to demonstrate that users were more engaged with a digital poster when their interaction was augmented by mid-air haptic feedback. In contrast to these predominantly hedonic added values, less consensus exists on whether mid-air haptics also adds to pragmatic aspects of the UX, such as perceived workload. Freeman et al. [12] did find that tactile feedback can enhance above-device gesture interactions with a smartphone (i.e. a more utilitarian task) but detected no preference for mid-air haptic feedback over vibrotactile feedback in this regard. Harrington et al. [6] reported a significant increase in accuracy for slider-bar tasks in a driving simulator when mid-air haptics was added, but Sand et al. [13] in turn did not find a similar effect for gesture-based button selection in VR.

Even though all these studies have provided new and enriching insights into the different aspects of the UX of mid-air haptics, none of them have considered the effects of prolonged use. By assessing the UX over a longer period of time, it becomes possible to investigate whether and how the UX of mid-air haptics would change over time.

2.2 Temporal Aspects of UX

The studies mentioned above demonstrate that in particular with regard to mid-air haptics, little attention has been paid to studying prolonged use of this new technology. However, a growing body of UX research on other product categories or technologies considers temporal aspects and their influence on how users' experience of

products or services changes over time. McCarthy and Wright [14] conceptualize experience with technology as consisting of four threads: the sensual, emotional, compositional and spatio-temporal (p. 80). The latter refers to a sense of space and time while using technology, showing the importance of time as an integral aspect of our experience with technology. Further on, McCarthy and Wright propose a tool for analysing how people make sense of technology introducing six processes of sense-making, which can occur at various moments in the use of technology and that can be analysed from the perspective of each of the threads. The six processes include anticipating use, connecting with a product or service, interpreting an unfolding experience, reflecting on the experience, appropriating an experience and recounting it to others. While not per se linear or in the order as presented here, the authors highlight how the various processes might differ between initial use or prolonged use.

While research on most of these processes is a more recent phenomenon, a lot of earlier scholars had already focused on users' initial intention to use. The original Technology Acceptance Model (TAM) considers perceived usefulness as the main predictor of intention to use, along with perceived ease of use [15]. Some later theoretical developments, which were based on the TAM, focused on technology acceptance in the consumer context, as opposed to the workplace, and added non-pragmatic UX components as well. An example is the Unified Theory of Technology Acceptance and Use 2 (UTAUT 2), where hedonic motivation or perceived enjoyment, defined as 'the fun or pleasure derived from using a technology' was added next to the pragmatic UX components [16]. In line with these theoretical models, Köse, Morschheuser, Hamari [17] found that if a product is perceived as mostly utilitarian, pragmatic aspects of the UX are dominant in predicting intention to use and continued use. However, when a product has a mostly hedonic nature, perceived usefulness loses its predictive power in favour of perceived enjoyment, which then becomes the main predictor of intention to use and prolonged use [17, 18].

In the meantime, some researchers have tried to capture various aspects of long-term use. Von Wilamowitz Moellendorff et al. [19] argue that our perception of the qualities of a product are dynamic and changeable over time. Their research on mobile phone use showed that as we get accustomed to a product, we develop and attach different weights to different qualities: whereas the initial focus might be on usability, this could shift to, e.g. novel functionality or communication of a favourable identity. Continuing the same line of thought, and based on two longitudinal user studies, Karapanos presents a framework of UX over time [20]. He shows how users initially evaluate a product based on its use, and that pragmatic quality, i.e. usefulness and ease-of-use, is of most importance in the beginning. However, after prolonged use, they evaluate the product based on their ownership of it and the importance of how well they identify with the product, i.e. what the product expressed about their self-identity in social contexts, increases. Furthermore, they found that the extent to which a product is found 'stimulating' (i.e. original, creative, new, innovative) has an effect on how beautiful it is considered to be, but this effect of 'stimulation' seems to diminish and makes place for 'identification' as the most important predictor of how beautiful a product is perceived.

2.3 Novelty Effect of Mid-Air Haptics

Related to the 'stimulation' mentioned above, scholars have identified an additional phenomenon to consider when assessing a product's UX; the so-called novelty effect, which often occurs at the first interaction(s) with a new technology (cf. 'connecting' in terms of McCarthy and Wright [14]). Novelty effect is defined by Koch et al. as 'an increased motivation to use something, or an increase in the perceived usability of something, on account of its newness' [21]). Koch also showed, however, that the usage patterns and/or perceived usability changes when novelty eventually fades [21].

As for mid-air haptic technology, most studies on its UX are cross-sectional and do not address the potential impact of a novelty effect [5, 6, 9, 12, 13]. In one study, a familiarization phase was included before starting with the experimental task, with the aim of mitigating a novelty effect. However, no assessment of perceived novelty was included to verify whether it actually decreased after the familiarization phase [11]. In a recent study [22], the impact of a novelty effect on the user experience of mid-air haptic feedback was tested by statistically controlling for perceived novelty. In this study, mid-air haptic feedback showed to provide added value on top of visual feedback in a gesture-based interface when considering attractiveness and pleasure during the interaction. However, these effects disappeared after statistically controlling for perceived novelty. This could imply that a decrease in novelty might go hand in hand with a decrease in attractiveness and pleasure. Longitudinal research is needed to investigate how the user experience of mid-air haptics evolves over time, and what happens when the novelty effect fades.

We are aware of only one study testing the repeated experience of mid-air haptics sensations: Ablart et al. [4] investigated the added value of mid-air haptics while watching one-minute movies at two points in time, with a time lapse of two weeks. They observed that mid-air haptic sensations increased the arousal ratings at both points in time, but that the skin conductance response (SCR) dropped at the second assessment, which reflects a drop in implicit arousal. This discrepancy between mid-air haptics' impact on self-rated arousal and implicit arousal over time is interesting from the perspective of the novelty effect. The results might be understood in terms of a fading novelty effect, reflected in the lower SCR. If this would be indeed the case, the unchanged self-rated arousal at both sessions could mean that subjective arousal is not particularly sensitive to a novelty effect.

3 The Study: 'Mixed-Method', 'Longitudinal' and 'in the Lab'

The present study evaluated the user experience of mid-air haptics over an extended period of time. In the following section, we will go over (a) the device and interface that were used for our participants to experience mid-air haptics; (b) participants and recruitment procedure; (c) the study set-up and procedure; (d) the UX assessment;

and (e) an excerpt of the main results. In the discussion section, we will further reflect, where relevant, on the decisions that were made against the light of our research purpose, as well as the implications and discerned best practices.

3.1 Apparatus

One company that began commercializing mid-air haptic technology is UltraLeap. Known before as UltraHaptics (before merging with LeapMotion), this company started as a spinoff from the university of Bristol (UK) and has developed a range of ultrasound mid-air haptic devices. As mentioned, the newness, cost and complexity of this technology make it not something that can be taken home or easily integrated in everyday devices and, as such, in the everyday life of participants. For our study, we used a Stratos Development kit by UltraLeap and linked it to a gesture-controlled home automation interface. It showed a groundplan of a house on which one could select and deselect rooms to then adjust the lighting, temperature, blinds and air conditioning through a set of four simple gestures (as shown in Figs. 1 and 2).

3.2 Participants

Participants were recruited on-site at the campus where the study took place. In total, 126 people signed up of which 31 were selected. We were very strict in the requirements for study participation in terms of availability and commitment: via an anonymized doodle, participants had to select eight time slots spread over a period of five weeks. For the first week, one long session (45 min) had to be selected. Over the second, third and fourth weeks six short sessions (15 min) had to be selected, and for the fifth week a final long session (45 min) again. Only when eight sessions were selected and distributed evenly over the five-week span, participants were considered

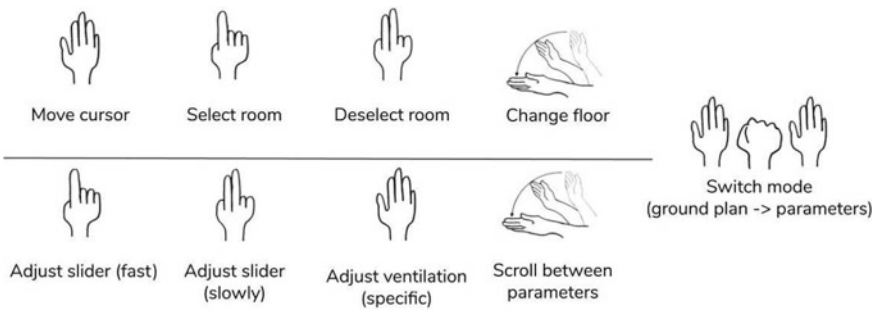

Fig. 1 Gestures to interact with the home automation interface

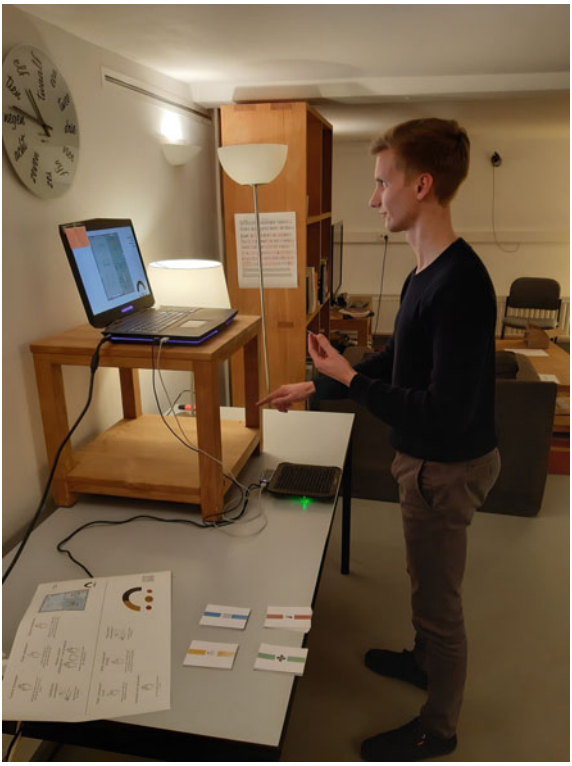

Fig. 2 Study set-up: participants used their right hand for gestural control above the UltraLeap kit and their left hand to draw and read scenario cards (cfr. 3.3 Procedure)

eligible. By having each participant define, in advance, their own dedicated five-week participation schedule, we hoped to decrease the chances of study dropout. Our strategy was successful as only one participant dropped out during the course of the study. One other participant encountered technical errors during the first session, resulting in valid data only in the final/closing session. Finally, we had to exclude one participant from the study because of an insufficient knowledge of Dutch, which was the language in which the experimental tasks, questionnaires and interviews were set up. As such, the initial number of 31 reduced to 28 final participants, which we considered as an acceptable number given the required engagement. Of the 28 remaining participants, the mean age was 20.79 (SD = 2.44), with an age range between 18 and 26. Six participants were male, and 22 were female. This study was approved by the local social and societal ethics committee: G- 2019 10 1780.

3.3 Procedure

Each participant came to our lab on campus at eight separate times. In each session, they were given a set of eight tasks to complete in the home automation system's interface described above by using the appropriate gestures. Gestures were either complemented by mid-air haptic feedback (first condition) or unadorned (second condition). These two conditions were presented in counterbalanced order over all sessions. As such, within each condition, participants each time completed four different tasks. These tasks were part of a narrative in which participants were members of an imaginary household and would encounter home-specific scenarios that required them to perform small tasks. An example of a scenario would be 'You have visitors tonight. Set the temperature of the living and dining room to a comfortable degree (21 °C). Then deselect the rooms again'. Scenarios were presented on cards randomly drawn from a face-down deck. There were four separate decks, each one corresponding to a different household functionality. The first deck contained tasks to adjust the lighting, the second deck had tasks to change the temperature (thermostat), the third one was related to the air conditioning and the last one to the window blinds. Two experimenters were allocated randomly to all sessions and each time followed the same experimental script. This means that participants were tested by two different experimenters across the eight sessions, based on random allocation.

3.3.1 Introductory Session

Upon arriving at our lab, participants first received extensive information about the five-week study schedule and procedure and were given the time to carefully read and sign the informed consent. After signing, participants were introduced to the mid-air haptic device and could familiarize with it through a range of sensations and patterns. Next, the home automation system was introduced. Participants were given time to get acquainted with the different gestures (Fig. 1). When they indicated feeling comfortable using the gestures, we asked what they expected from the combination of mid-air haptic feedback with this gestural home automation interface. The answers to this question were audio recorded and visited again during the last session. At this point, the actual experimental tasks started. Participants started interacting with the home automation system either with or without mid-air haptic feedback, depending on the order they were assigned to (counterbalanced). In each condition, they picked one scenario card from each of the four decks and completed the tasks one by one. We emphasized that they had to perform the tasks as accurately as possible but not as fast as possible. It was important that they did not feel stressed or hurried while interacting with the home automation system, but rather calm in order to be able to experience the interaction to the fullest. After completing four scenarios in the first condition (either with or without mid-air haptic feedback), they received a questionnaire to assess their experience with the home automation system (cfr. UX Assessment, Questionnaire section). They then repeated this with four new scenario

cards (with or without mid-air haptics, depending on the previous condition) and filled out the same questionnaire for a second time. When both conditions were completed, a qualitative segment followed in which participants answered a set of open-ended questions (cfr. UX Assessment, Interview). At the end of the session, which took about 45 min, participants received a €10 bol.com voucher.

3.3.2 Repeated Sessions

After the more extensive introductory session, participants returned for six short repeated sessions. In these sessions, participants again completed four tasks in each condition (with/without mid-air haptic sensations) in counterbalanced order. These sessions took about 15 min to complete. Again, in each condition, one scenario card was picked from each of the four decks, totalling four different tasks in each condition, and eight different tasks in total. To enable participants to quickly drop by and perform the sessions right before, after, or in between classes, no questionnaire or interview was included in these short sessions.

3.3.3 Closing Session

The last session was again a long one and took about 40 to 45 min to finish. After completing the scenarios (picked from the four different card decks), participants received the same questionnaire as in the introductory session for each condition. When both conditions were completed, we revisited the expectations they expressed in session 1 and asked them whether they were met or not. This question was used to instigate the conversation on their overall experience. In addition, it gave participants the chance to nuance their questionnaire answers further. As such, we again elaborated on the same variables as those from the first session, to broadly understand why certain experiences changed or did not change over the five-week period (see also Sects. 3.4 and 3.5). At the end of this session, participants received the €40 bol.com voucher, as a reward for participating in all eight sessions.

3.4 UX Assessment

To (a) evaluate the changes in user experience over time quantitatively; and (b) gain deeper insights in participants' perception of the experience using a qualitative approach, we applied a mixed-method design that offered a broad understanding of our participants' attitude towards mid-air haptics. Here, we describe both the questionnaires that were used, as well as how this data was enriched with insights captured by the open-ended interviews. We then briefly discuss some of the main results.

3.4.1 Questionnaires

In line with the models introduced in the Related Work section, our questionnaire assessed both pragmatic and hedonic aspects of the UX, as well as the valence and arousal of the user's emotional reactions.

The questionnaire started with some general questions on age, gender and handedness. Subsequently, participants had to indicate the condition they had just completed tasks in: with or without mid-air haptic feedback. This was intended as an exclusion criterion item: it enabled us to filter out participants who had not paid any attention to the presence or absence of mid-air haptic feedback. Subsequently, a combination of existing standardized questionnaires was included in randomized order: the Affective Slider (AS) [23], User Experience Questionnaire (UEQ) [24], User Engagement Scale Short-Form (UES-SF) [25], perceived usefulness and perceived ease of use of the TAM [15, 18], enjoyment, continued use and user conception based on Köse et al. [17] and Van der Heijden [18]. We thus obtained four completed questionnaires from each participant: two on their experience with the interface with mid-air haptic sensations (one from the first session and a second one from the last session) and two about their experience without mid-air haptic sensations, again from both the first and the last session.

3.5 Interview

In addition to the more standardized quantitative assessments described above, both our first and last sessions ended with a set of open-ended questions which were audio recorded and transcribed for thematic analysis. The purpose of this set of wrap-up questions was to allow participants to elaborate further on their quantitative responses, stimulating them to reflect and add nuance and supplementary information that was not recorded in the questionnaire. The interview questions were therefore mapped to segments from the quantitative questionnaire: we asked about 'efficiency' (i.e. 'did you find the home automation interface more efficient with or without the mid-air haptic feedback?'), 'ease of use' (i.e. 'did you find the home automation interface easier to use with or without the mid-air haptic feedback?') and in the same fashion 'enjoyment' and 'continued use' (i.e. 'would the addition of mid-air haptic to the interface have an influence on whether you continue using it?'). In addition, we also asked about their overall preference (with or without mid-air haptic feedback).

3.6 Results Excerpt

To illustrate how the questionnaire data were analysed, we report in this section the statistical analyses used to obtain the results of three variables of main interest: enjoyment (hedonic UX), ease of use (pragmatic UX) and the valence of the

emotional reaction. 'Enjoyment' was assessed as the mean score on four 7-point Likert items [17, 18]. Similarly, 'ease of use' was also assessed as the mean score on four 7-point Likert items [15, 18]. The valence of the emotional reaction was assessed using the affective slider, with a scale from 0 to 100 [23]. We used repeated measures (RM) ANOVA (R package ez [26]), with two within-participant factors: condition (with/without mid-air haptics) and session (session 1/session 8). A separate RM ANOVA was performed for the three dependent variables: enjoyment, ease of use and valence. We tested for both the main effects of condition and time and the interaction between both. All analyses were performed on complete data of 28 participants.

With enjoyment as dependent variable, we observed a main effect of condition, $F(1, 27) = 8.97$, $p < 0.01$, $\eta_G^2 = 0.02$, with the condition including mid-air haptics leading to significantly higher enjoyment than the condition without mid-air haptics. This can be understood as an added value of mid-air haptics in terms of enjoyment. A main effect of session was observed as well, $F(1, 27) = 15.10$, $p < 0.001$, $\eta_G^2 = 0.08$, with significantly lower enjoyment during the last session compared to the first session. Although there appears to be a decrease in the added value of mid-air haptics regarding enjoyment, when comparing session 1 with session 8 (see Fig. 3), this interaction effect was not statistically significant, $F(1, 27) = 1.18$, $p = 0.29$, $\eta_G^2 = 0.00$, which means that the added value of mid-air haptics in terms of enjoyment was similar during the first and last session.

Regarding ease of use, only a significant main effect of session was observed, $F(1, 27) = 7.32$, $p = 0.01$, $\eta_G^2 = 0.06$, with overall significantly higher ease of use during the last, compared to the first session. Condition showed no main effect, $F(1, 27) = 0.04$, $p = 0.85$, $\eta_G^2 = 0.00$, which means that there was no added value

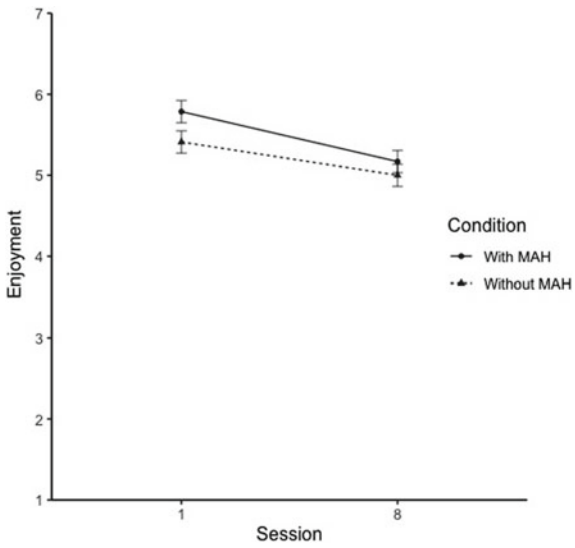

Fig. 3 A main effect of condition and session, but no interaction effect, when considering enjoyment

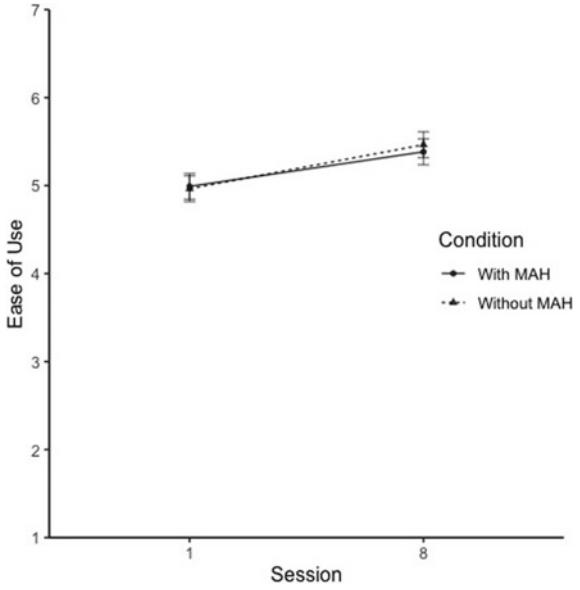

Fig. 4 A main effect of session, but no main effect of condition and no interaction effect, when considering ease of use

of mid-air haptics in terms of ease of use. Finally, no interaction effect between condition and session was present either, $F(1, 27) = 0.28$, $p = 0.60$, $\eta_G^2 = 0.00$ (see Fig. 4).

Concerning the valence of participants' emotional reaction: there was a main effect of condition, $F(1, 27) = 8.11$, $p < 0.01$, $\eta_G^2 = 0.05$, a main effect of session, $F(1, 27) = 7.61$, $p = 0.01$, $\eta_G^2 = 0.04$, and a significant interaction effect between condition and session, $F(1, 27) = 4.82$, $p < 0.05$, $\eta_G^2 = 0.02$. Therefore, we only interpreted this interaction effect (see Fig. 5). At session 1, the home automation system with mid-air haptics led to more experienced pleasure than without mid-air haptics, but this added value of mid-air haptics disappeared at session 8. This means that after repeated use, the added value of mid-air haptics in terms of experienced pleasure disappeared.

This is a clear illustration of what could be considered a novelty effect: initially, there was a significant increase in experienced pleasure due to the newness of the mid-air haptic sensations, but this effect disappeared after repeated use, when the novelty possibly faded away. When considering enjoyment, there appears to be a similar trend towards a fading added value of mid-air haptics at session 8 (Fig. 3), however this interaction effect was not statistically significant. Concerning ease use, we observed no evidence for a novelty effect, as the presence of mid-air haptic sensations did not have any impact at all on ease of use (Fig. 4).

The qualitative segment uncovered how nuanced and ambiguous preferences and UX experiences actually were. Only half of the participants maintained their initial preference (pro, contra or indifferent of mid-air haptics). Participants who after all eight sessions retained their preference for actuation of the home automation interface

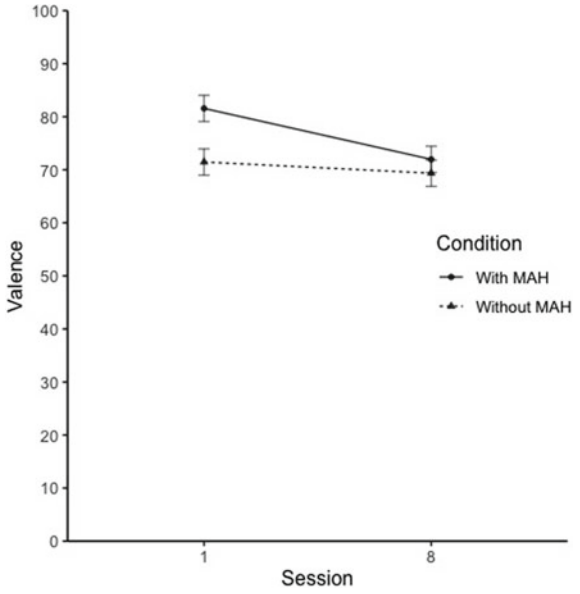

Fig. 5 A main effect of condition and session, and a significant interaction effect between both, when considering valence of the emotional reaction

with mid-air haptics often mentioned how the mid-air haptics made the interaction more pleasant but not necessarily more practical. Some of them mentioned this was due to the guidance and confirmation they received through the mid-air haptic sensations, while others associated it with a heightened sense of agency over the system. In contrast, participants who preferred the interface without mid-air haptics very often mentioned a sense of being startled by the sudden sensations, making the experience unpredictable and thereby uneasy and unpleasant. Additionally, an interesting temporal component that was shown by the interviews was how mid-air haptics were either preferred in the beginning sessions versus only later on. Some participants described how the mid-air haptics helped them to get acquainted with the home automation system and the gestures used to control it, while others experienced the mid-air haptics as distracting at first, but grew fond of them once they got used to them.

4 Reflections and Implications for Longitudinal Research

As longitudinal research in HCI is rather scarce, especially in an in-lab setting, we now share the main methodological and practical take-outs of the present study, as well as reflect on the decisions (and their implications) that were made.

4.1 Planning

For in-lab studies, participants will engage in dedicated, stand-alone interaction sessions for which they need to visit the research lab on a regular basis. Especially when the amount of participants as well as the amount of sessions is high (as was the case in our study), it is paramount to plan well in advance. This facilitates the process both for participants as well as the researcher. That is why we had the people who showed interest to participate (n = 126) fill out time slots on an anonymized Doodle calendar with the chosen time slots disappearing as options for new participants. Everyone was instructed to spread their sessions evenly over the five-week period. The main reason for this was that we wanted participants to have a well divided interaction interval with the interface (rather than, e.g. seven interactions during the first week and only five weeks later an eight time). In addition, this not only made the effort for participants themselves more dispersed, but also allowed us as researchers to maintain a balanced data collection schedule. Unfortunately, having a predefined schedule is no guarantee for each session to take place exactly as planned. Technical issues, session cancellations or requests for rescheduling unavoidably take place when having over a hundred test moments. As such, it is important to build in buffer time as well as flexibility. In our case, we wanted catch-up sessions to take place as close as possible to the original test moment, for an even spread of sessions over time to be maintained for each participant. Participants were made aware that they would only receive the last part of their remuneration (cfr. infra) after attending all eight sessions.

In conclusion, having participants select their own time slots not only allows researchers to keep a structured overview, but also, and maybe even more importantly, implies to the participants that they have committed to a schedule of their liking. As such, there is less excuse to not show up or to drop out. Nonetheless, rescheduling will happen, and it is of importance to be both prepared as well as agile in this regard. Even though these measures might seem self-evident, we want to emphasize that this could make a difference in terms of retention and study dropout.

4.2 Trade-Off on Session Duration

Next to the session frequency, the session duration is of importance too when it comes to participant attrition. By keeping sessions short, as we did for the largest part of our study, participants will tend to remain more motivated and come back for each session. With relatively long sessions at the start and the end of the study, we thus deliberately decided to keep the intermediate sessions short and not collect data in them. This decision came at a cost, however. Having only two points of data collection for each participant causes restraint on the conclusions that can be drawn from a longitudinal perspective. The current data allows us to report on the changes in enjoyment, ease of use and valence between the start and finish of the study, but not

on possible shifts in between. One might wonder, for example, whether valence might have increased in session 2 and 3, to only then make the reported drop. The'shape of change' for our variables is in that sense unknown, and asks for further research. When mid-air haptic technology becomes more easily integrable in daily devices and thereby more widely available, the possibilities to have more data collection points will obviously increase too. Capturing data at several points in time over a prolonged period feasibility of collecting data entries at multiple points in time would enrich our understanding of longterm UX of mid-air haptics by shedding light on this currently occluded period between the first and last session. As such, it might provide answers to questions such as when the decline in enjoyment and valence exactly set in, and whether variables first show trends in other directions.

4.3 Recruitment

It is self-explanatory that a big, heterogeneous and representative participant pool is desirable in order to be able to generalize results as much as possible. However, there are always practical limitations to consider, especially in the case of a longitudinal study. As a participant, committing to a longitudinal in-lab study not only requires a considerable amount of time but also demands repeated logistic efforts. If participants are required to come to the lab often and frequently, it is sensible to take into account geographical factors. For our study, we therefore deliberately recruited at the campus of our lab, to reach potential participants who were there regularly and (hopefully) lived nearby. In addition to practical reasons, this approach had, for our study in particular, the additional advantage of reaching the target group that is known to be most perceptible to mid-air haptics. Research has shown that the sensitivity to, and ability of perceiving mid-air haptic sensations, declines with age [27].

4.4 Fun Factor

Data collection can be a tedious process. For the participant, there is often not much variation and tasks tend to get monotone and boring. Especially for multiple sessions with repeated tasks, you might consider gamifying the process or add a narrative to it. We included a simple story of participants being household members who, in their homes, encountered everyday 'scenarios'. These scenarios were presented through the card decks described previously. What was actually a very plain on/off exposure to mid-air haptic feedback now became a set of micro narratives that gave purpose to the interactions. When conducting a longitudinal study with repeated contact moments, we suggest knitting these stand-alone sessions into a bigger whole, possibly with a conclusion to be reached at the end.

4.5 Remuneration

If the study's budget allows remuneration for participants, this is of course an obvious benefit in terms of recruitment and retention. However, there are multiple ways to approach its distribution and spread. First of all, in some studies there is only one or a few 'rewards' or 'prizes' distributed among participants. Especially when asking for a long-term commitment, this might feel insufficient and unfair towards participants who miss out, which is why we would recommend distributing the budget evenly and compensating each participant with at least a small, even, remuneration. In addition, instead of foreseeing this compensation at the end of the study, one might consider spreading it. By doing so, retention and loyalty can be encouraged implicitly. The budget of our study allowed us to thank our participants with a €50 gift voucher each. This in itself is of course a significant amount and will facilitate retention. However, we did decide to give €10 already after the first session, and the remaining €40 at the end of the last session. This metaphorically 'reeled in' participants for the initial session, to then have them 'bought in" sufficiently to last until the final one.

5 Conclusion and Future Research

This chapter describes the set-up of a longitudinal in-lab study of which the goal was to assess the UX of a novel technology—ultrasound mid-air haptic feedback—over a prolonged period of time. We wanted to investigate how the hedonic and pragmatic UX of mid-air haptics would evolve over time and how participants' emotional reactions to this type of feedback would change. Building on a set of existing models and questionnaires, we applied a mixed-method design to generate a broad understanding of our participants' attitude towards mid-air haptics. During eight sessions (spread over a five-week period) participants interacted with a gesture-controlled home automation system, augmented with mid-air haptic feedback half of the time. This approach provided unprecedented insights and understandings of how people experience mid-air haptic technology, in particular after repeated use. Results showed that the pleasure and enjoyment participants experienced after their first interactions with mid-air haptics, significantly decreased over time. Regarding experienced pleasure, the added value of mid-air haptics at session 1 even disappeared at session 8, indicating that after repeated use, the presence of mid-air haptics no longer led to higher experienced pleasure compared to when absent. Although a similar trend for the hedonic UX of mid-air haptics (enjoyment) appeared to be present, this was not statistically significant. This means that the added value of mid-air haptics regarding enjoyment was relatively stable across both sessions, with a general decrease in enjoyment (whether or not mid-air haptics were present) from session 1 to session 8. Interestingly, participants reported no added value of mid-air haptics with regard to ease of use (an aspect of the pragmatic UX). Overall (whether or not mid-air haptics were present), the ease of use significantly increased from

session 1 to session 8, as can be expected when interacting regularly with a new device. The findings from the interviews confirmed these results for the largest part, but at the same time revealed much nuance and ambiguity in participants' preferences. Hedonic and pragmatic preferences were regularly in conflict with each other and for some participants the added value of the mid-air haptic feedback increased instead of decreased. They, for example, mentioned that it took some time getting used to the sensations, but that the 'startling effect' over time faded, accustoming them to the mid-air haptics. Based on the participant pool we had, we could not discern interpersonal traits that provided a salient account or explanation for this. Future work to assess the relation between such personal traits and a negative vs. positive aptitude for mid-air haptic feedback will be interesting and necessary in that regard. Another remark to be made here is that measuring the UX of mid-air haptics typically happens with the mid-air haptic sensations being part of a larger interface, in our case: a home automation system. Earlier research (e.g. [6, 11]) similarly assessed the mid-air haptics as part of a larger whole, seeing that mid-air haptics as a stand-alone output typically makes less sense and lacks applicability and relevance (for exceptions, see Van den Bogart et al., 2019). Although the interface in casu (and its either pragmatic or hedonic character) unavoidably influences the user experience in its entirety to some extent, we mitigated this as much as possible by calculating the difference score between participant's evaluation of the home automation system with versus without mid-air haptics.

As mentioned, given the novelty and uncommonness of mid-air haptic technology, it was not possible for participants to interact with it on an individual day-to-day basis in their own home. A lab-setting was needed for our study. In addition to the traditional challenges of longitudinal research, this confinement to the lab brought about extra challenges, mainly in terms of participant engagement and retention. We therefore deliberately stuck to two data collection points in order not to overburden participants. The trade-off for this decision, however, was that we can only report on these measures and not on what happened to our variables in between. This leaves other questions (e.g. 'is the decline of valence and enjoyment linear?') unanswered and up to future research.

Regardless, we have applied and discussed additional techniques (other than financial remuneration) to foster participant retention and avoid study dropout. By gamifying or adding a narrative to required study tasks, their obligatory character can be dissolved, making continued participation more pleasant and casual. A well-planned schedule and time table, created by the participants themselves, not only keeps things clear and structured, but also increases the participants' sense of commitment. In addition, we recommend considering pragmatic elements when recruiting participants in order for the process and logistics to remain feasible for both them and the researchers.

References

1. Carter T, Seah SA, Long B, Drinkwater B, Subramanian S (2013) UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. In: Proceedings of the 26th Annual ACM symposium on user interface software and technology, pp 505–514. UIST '13, ACM, New York, NY, USA. https://doi.org/10.1145/2501988.2502018
2. Wilson G, Carter T, Subramanian S, Brewster SA (2014) Perception of ultrasonic haptic feedback on the hand: localisation and apparent motion. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1133–1142. CHI '14, ACM, New York, NY, USA. https://doi.org/10.1145/2556288.2557033
3. Obrist M, Seah SA, Subramanian S (2013) Talking about tactile experiences. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1659–1668. CHI '13, ACM, New York, NY, USA. https://doi.org/10.1145/2470654.2466220
4. Ablart D, Velasco C, Obrist M (2017) Integrating mid-air haptics into movie experiences. In: Proceedings of the 2017 ACM international conference on interactive experiences for TV and online video. pp 77–84. TVX '17, Association for Computing Machinery, Hilversum, The Netherlands. https://doi.org/10.1145/3077548.3077551
5. Hwang I, Son H, Kim JR (2017) AirPiano: enhancing music playing experience in virtual reality with mid-air haptic feedback. In: 2017 IEEE world haptics conference (WHC), pp 213–218. https://doi.org/10.1109/WHC.2017.7989903
6. Harrington K, Large DR, Burnett G, Georgiou O (2018) Exploring the use of mid-air ultrasonic feedback to enhance automotive user interfaces. In: Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications, pp 11–20. AutomotiveUI '18, Association for Computing Machinery, Toronto, ON, Canada. https://doi.org/10.1145/3239060.3239089
7. Hassenzahl M (2004) The interplay of beauty, goodness, and usability in interactive products. Hum-Comput Interact 19(4):319–349. https://doi.org/10.1207/s15327051hci1904_2
8. Mahlke S, Thüring M 2007) Studying antecedents of emotional experiences in interactive contexts. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 915–918. CHI '07, Association for Computing Machinery, San Jose, California, USA. https://doi.org/10.1145/1240624.1240762
9. Maggioni E, Agostinelli E, Obrist M (2017) Measuring the added value of haptic feedback. In: 2017 ninth international conference on quality of multimedia experience (QoMEX), pp 1–6. https://doi.org/10.1109/QoMEX.2017.7965670, ISSN: 2472-7814
10. Kujala S, Mugge R, Miron-Shatz T (Feb 2017) The role of expectations in service evaluation: a longitudinal study of a proximity mobile payment service. Int J Hum-Comput Stud 98:51–61. https://doi.org/10.1016/j.ijhcs.2016.09.011, http://www.sciencedirect.com/science/article/pii/S1071581916301239
11. Limerick H, Hayden R, Beattie D, Georgiou O, Müller J (2019) User engagement for mid-air haptic interactions with digital signage. In: Proceedings of the 8th ACM international symposium on pervasive displays, pp 1–7. PerDis '19, Association for Computing Machinery, Palermo, Italy. https://doi.org/10.1145/3321335.3324944
12. Freeman E, Brewster S, Lantz V (2014) Tactile feedback for above-device gesture interfaces: adding touch to touchless interactions. In: Proceedings of the 16th international conference on multimodal interaction—ICMI '14. ACM Press, Istanbul, pp 419–426. https://doi.org/10.1145/2663204.2663280
13. Sand A, Rakkolainen I, Isokoski P, Kangas J, Raisamo R, Palovuori K (2015) Head-mounted display with mid-air tactile feedback. In: Proceedings of the 21st ACM symposium on virtual reality software and technology, pp 51–58. VRST '15, ACM, New York, NY, USA. https://doi.org/10.1145/2821592.2821593
14. McCarthy J, Wright P (2004) Technology as experience. https://doi.org/10.1145/1015530.1015549
15. Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 13(3):319–340 (1989). https://doi.org/10.2307/249008

16. Venkatesh V, Thong JYL, Xu X (2012) Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 36(1):157–178. https://doi.org/10.2307/41410412
17. Köse DB, Morschheuser B, Hamari J (2019) Is it a tool or a toy? How user's conception of a system's purpose affects their experience and use. Int J Inform Manage 49:461–474. https://doi.org/10.1016/j.ijinfomgt.2019.07.016, http://www.sciencedirect.com/science/article/pii/S0268401219301550
18. van der Heijden H (2004) User acceptance of hedonic information systems. MIS Q 28(4):695–704. https://doi.org/10.2307/25148660
19. Wilamowitz Moellendorff M, Hassenzahl M, Platz A (2006) Dynamics of user experience: How the perceived quality of mobile phones changes over time. In: User experience—towards a unified view, workshop at the 4th Nordic conference on human-computer interaction
20. Karapanos E (2013) User experience over time. In: Karapanos E (ed) Modeling users' experiences with interactive systems. Studies in Computational Intelligence, Springer, Berlin, Heidelberg, pp 57–83. https://doi.org/10.1007/978-3-642-31000-3_4
21. Koch M, von Luck K, Schwarzer J, Draheim S (2018) The novelty effect in large display deployments—experiences and lessons-learned for evaluating prototypes. https://doi.org/10.18420/ecscw2018_3, https://dl.eusset.eu/handle/20.500.12015/3115. (Accepted 2018-04-30T19:58:18Z)
22. Rutten I, Geerts D (2020) Better because it's new: the impact of perceived novelty on the added value of mid-air haptic feedback, p 13
23. Betella A, Verschure PFMJ (2016) The affective slider: a digital self-assessment scale for the measurement of human emotions. PLOS ONE 11(2)
24. Laugwitz B, Held T, Schrepp M (2008) Construction and evaluation of a user experience questionnaire. In: Holzinger A (ed) HCI and usability for education and work. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, pp 63–76. https://doi.org/10.1007/978-3-540-89350-9_6
25. O'Brien HL, Toms EG (2013) Examining the generalizability of the User Engagement Scale (UES) in exploratory search. Inform Process Manage 49(5):1092–1107. https://doi.org/10.1016/j.ipm.2012.08.005, https://linkinghub.elsevier.com/retrieve/pii/S0306457312001124
26. Lawrence MA (2016) ez: Easy analysis and visualization of factorial experiments. https://CRAN.R-project.org/package=ez
27. Rutten I, Frier W, Van den Bogaert L, Geerts D (2019) Invisible touch: how identifiable are mid-air haptic shapes? In: Extended abstracts of the 2019 CHI conference on human factors in computing systems—CHI EA '19. ACM Press, Glasgow, pp 1–6. https://doi.org/10.1145/3290607.3313004