UCLouvain

LOUVAIN
School of Management

Louvain School of Management

Analysis of User Experience of Interactive Applications by Analog Astronauts

Author:	Dylan Goffinet, Pierre Huljev
Supervisor:	Jean Vanderdonckt
Academic year:	2024-2025
Master [60]:	Management

Declaration Regarding AI Tool Usage in Master's Thesis

During the preparation of this master's thesis, the authors utilized DeepL and Scite for the following purpose:

1.DeepL: DeepL Translator and DeepL Write were utilised to translate, correct, and enhance the tone of several passages of this thesis.

2. Scite AI: was used to search sources about research questions and enrich the overall literature of the thesis.

3. After using DeepL and Scite, the authors diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

Through this declaration, we affirm that the content of this master's thesis reflects our original work, augmented by the responsible use of AI.

Acknowledgements

First, we would like to thank and express our gratitude to our Master's thesis supervisor, Professor Jean Vanderdonckt. Thank you for having confidence in us and giving us a subject that we think is close to your heart. Thank you also for accompanying us throughout the year and always being available and open to discussion.

We would also like to thank our family and close friends for their support throughout the year and for their efforts to understand our subject, which was not very successful.

We would also like to thank the Groupement des Cercles Louvanistes, the 7, the OCN, the AVAO, and all our friends from the master 60 who supported us throughout the year.

This page was intentionally left blank.

Abstract

This thesis is part of a wider study of Human-Computer Interactions (HCIs) in the context of the conquest of space. The capacity to adapt HCIs to extreme environments such as space is fundamental to our potential to colonise other planets. The aim of this study was to analyse the results of user experiments involving five different devices utilised as part of the M.A.R.S. UClouvain mission. The devices included Romie (an advanced modelling and scheduling management application), EZExfix (an external fixator for bone fractures), Tap Strap 2 (a single-handed, all-in-one wearable gesture controller), Harmony (an LLM designed to assist astronauts), and a Parrot drone. This thesis establishes the foundations for improving the adaptability and usability of interactive systems in hostile environments. To achieve this, data collection must be broadened to encompass both physiological metrics and user experience (UX), with the aim of investigating interactions between systems. Furthermore, HCI frameworks should be integrated into the product design process. Attention must also be paid to user diversity in order to create inclusive systems suitable for future space missions.

Contents

List of Figures

Chapter 1

Introduction

This thesis constitutes a component of a comprehensive project that investigates HCI within the paradigm of space exploration. The objective of this study is to analyse feedback from the MDRS research centre in the Utah desert as part of the M.A.R.S. UCLouvain project. These feedbacks will be analysed, and suggestions based on these analyses will be made in order to improve the user experience of a set of interactive systems.

The structure of this thesis is as follows. Chapter 2 offers a thorough exposition and delineation of the diverse interactive systems that have been examined and evaluated. In Chapter 3, an overview of the extant literature on human-computer interaction, user experience, and the design framework for extraterrestrial applicability is provided. In Chapter 4, the methods employed to analyse user experience and ascertain the presence of consistency across the different studies are presented. In Chapter 5, the results of the studies and the interpretation thereof are presented. The final chapter, Chapter 6, presents the study's conclusions, contributions, and managerial implications. The thesis's limitations are also highlighted, and recommendations for future research are provided.

Chapter 2

Studied devices

2.1 Romie

Romie is an advanced planning and scheduling (APS) tool that enables operations to be modelled, planned and optimised while taking account of temporal uncertainty (variable task duration) and multiple performance indicators (KPIs) Saint-Guillain et al., 2023.

Romie allows the user to model visually at any time (Saint-Guillain et al., 2023):

- The human and material resources.
- The tasks to be carried out.
- Time constraints.
- Uncertainties over task duration, via modified PERT-type distributions.

The implementation of these measures is driven by the objective of ensuring maximum accessibility for the entire population, thereby empowering non-specialists to engage in the creation of their own models. Furthermore, it is a versatile instrument that lends itself to application across a range of domains.

The stochastic approach of the `Romie` model has been demonstrated to maximise the probability of successful planning, limit the necessity for adjustments during execution, and minimise stress, delays and additional costs (Saint-Guillain et al., 2023).

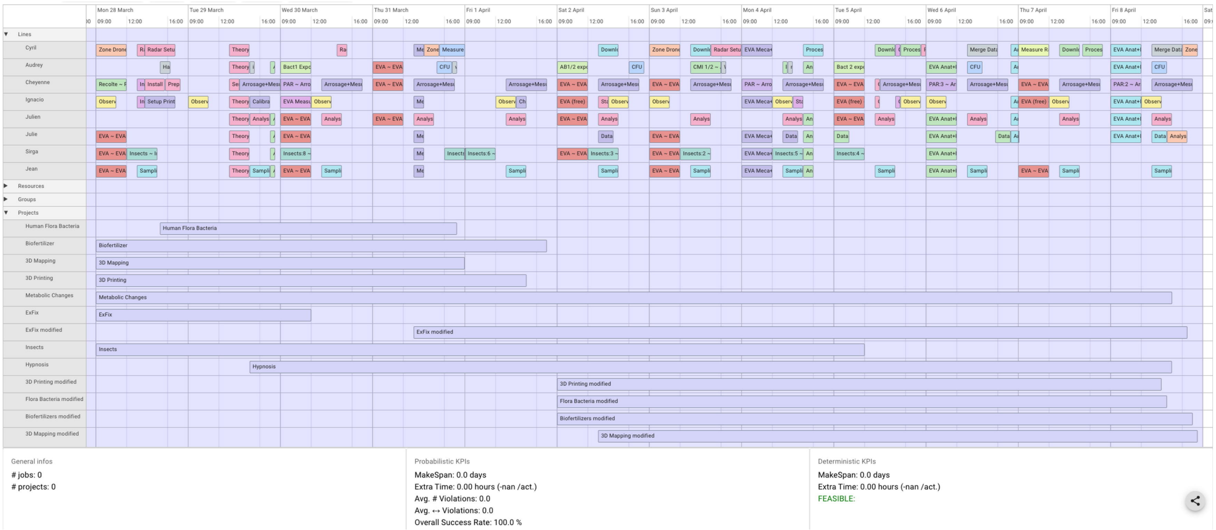


Figure 2.1: Final schedule at the end of the mission using Romie Saint-Guillain et al., 2023

2.2 EZExFix

Figure 2.2: Example of an artificial leg broken and fixed using an EZExFix (Manon et al., 2023)

The EZExFix is an external fixator that has been designed to enable non-surgical individuals to treat long bone fractures, thereby reducing dependence on surgical expertise. The application of the EZExFix has the potential to facilitate rapid and effective medical intervention in critical circumstances, thereby diminishing the necessity for conventional surgical proficiency. The EZExFix's compact and portable design offers significant advantages, rendering it as a promising solution for a variety of demanding environments, including space missions and emergency situations on Earth (Manon et al., 2023).

2.3 Tap Strap 2

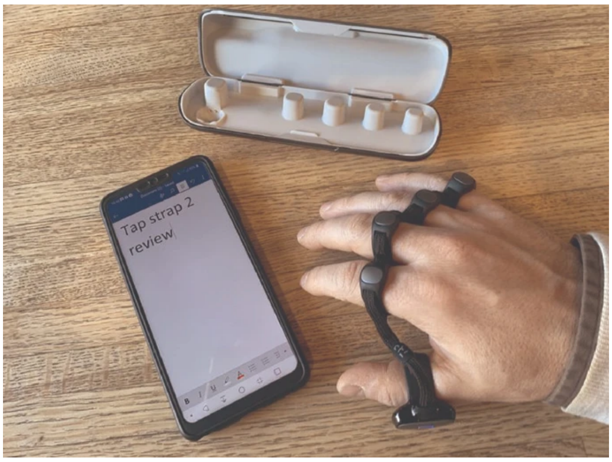


Figure 2.3: The Tap Strap 2 and its case (Mrazek et al., 2021)

Tap Strap 2 is a single-handed all-in-one wearable device that functions as a keyboard, mouse and air gesture controller. The device was designed for use in both gesture-based and touchless input, rendering it particularly relevant for use in confined and hostile environments. In such contexts, conventional input devices may be impractical due to the equipment required to conduct the missions. The objective of the wearable is to provide intuitive and versatile control for interaction with computer systems whilst minimising the need for surface-based interfaces (Mrazek et al., 2021).

Taps are detected through the use of built-in accelerometers in each ring. The data is then transmitted to and processed by a microcontroller unit (MCU) located within the thumb ring, which serves to determine the input (Mrazek et al., 2021).

Tap strap can be used in three modes (Mrazek et al., 2021):

- **Keyboard**: allows users to type characters by tapping combinations of fingers on a hard surface.
- **Mouse**: the thumb ring is equipped with an optical sensor that enables the user to control a computer mouse with their finger.
- **Air Gesture**: the device utilises a 6-axis inertial measurement unit (IMU) to detect hand movements in space. It allows to control a cursor at distance.

Tap Strap is then distinguished by its TapGenius learning application, which is reputed to be more intuitive than competing applications. It also features an integrated 3-in-1 mode and five light, open rings that allow the user to keep their hands partially free.

2.4 Parrot drone

Only one type of drone was used in the MARS UCLouvain studies. The Parrot ANAFI was utilized for the purpose of terrain mapping, in addition to monitoring the MDRS station infrastructure ("M.A.R.S. UCLouvain - Martian simulation", n.d.) and identifying potential entrances to subterranean structures that could serve as natural refuges against solar and cosmic radiation on Mars Tribolet and Officer, n.d.

Figure 2.4: Parrot ANAFI drone illustration ("Review: Hands On With the Parrot ANAFI - DRONELIFE", n.d.)

2.5 Harmony LLM

Large Language Models (LLMs) represent a significant advancement in the field of artificial intelligence, particularly in natural language processing (NLP). These models are built upon neural network architectures, primarily transformers, which excel at processing and generating human-like text. LLMs like GPT-3, GPT-4, and others are trained on expansive corpora of text, enabling them to capture intricate patterns and contexts inherent in human language (Karabacak and Margetis, 2023; Sarumi and Heider, 2024). By using LLMs for space exploration missions, it could enhance decision-making processes, simplify human-computer interactions and improve the accessibility of complex data for astronauts and mission control.

Chapter 3

Related works

3.1 Human-Computer Interaction

Human-Computer Interaction (HCI) is a multidisciplinary field of study. Its main purpose is to examine the interaction between humans and computer systems, as well as the extent to which computers have been developed to interact successfully with human beings Machno et al., 2015.

HCI was initially established as an academic discipline within the field of computer science. The primary objective of this initiative was to optimise productivity, with a particular emphasis on word processing software and data management tools. The present scope has now been expanded to encompass data visualisation, information technology systems and collaborative systems, among others Carroll, 2009.

3.1.1 HCI applied in space exploration

The concept of space travel and humanity's potential evolution into an interplanetary species have long been subjects of profound fascination for the human imagination. The pursuit of space exploration is driven by two objectives: advancing fundamental scientific knowledge and stimulating technological innovation. This research endeavour also fosters the emergence of new industries, thereby contributing to the economic and social development of humanity. However, maintaining human health, well-being and productivity in space is one of the most challenging aspects of contemporary space programmes. The human body evolved in the specific environment of Earth and is only able to survive the extreme conditions of space, such as high radiation and low gravity, with difficulty. Consequently, researchers have developed a variety of human-computer interfaces to enhance physical and mental performance in space. Recent advancements in aerospace engineering, coupled with democratised access to space, have made space research increasingly viable and accessible. Consequently, researchers in HCI are presented with a significant opportunity to contribute to the ambitious undertaking of space exploration. This could involve designing new types of interactive systems and computer interfaces capable of supporting people in their daily activities, both on Earth and in space and elsewhere in the solar system (Pataranutaporn et al., 2021).

By looking at the historical interplay between HCI and space exploration reveals a mutual influence and interconnection. For example, key concepts in HCI were first developed in the context of space exploration. These concepts were conceived in the objective to let computers take over human tasks and support astronauts in space (Pataranutaporn et al., 2022).

One example of this is virtual reality (VR) technology, which was originally developed by HCI researchers at NASA to train astronauts. The pioneering work of Margaret Hamilton and her team on algorithms and interfaces for Apollo 11 in 1969 can also be examined. The system was designed around the principles of robustness and human-centred quality. Consequently, it facilitated trust and collaboration between the astronauts and the computer, as well as enabling the interpretation of error messages and the delegation of low-priority tasks to the computer. This ultimately contributed to the successful completion of the mission to the Moon and back (Pataranutaporn et al., 2022).

This illustrates the close connection between HCI and space exploration, leading to the development of new space missions, scientific discoveries, and technologies that benefit people on Earth and in space. Consequently, advances in HCI in space would also impact people on Earth (Pataranutaporn et al., 2022).

3.1.2 SpaceCHI: Advancing Human-Computer Interaction Systems for Space Exploration

SpaceCHI is an initiative of the HCI scientific community, aiming to support human physical and mental performance in extraterrestrial environments Pataranutaporn et al., 2022. The main goal is to design new interactive computer interfaces to support humans living and working in space. SpaceCHI focuses on the areas of space exploration that require HCI knowledge and expertise. Examples include exoskeletons, virtual and augmented reality systems, and zero-gravity musical interfaces for entertainment. Vanderdonckt et al., 2024.

Contributions to SpaceCHI are diverse and substantial. These include tools for crew collaboration and mission planning, the resilience of human systems for deep space missions, human–robot interaction in extraterrestrial missions, the food experience in space travel, and the influence of extraterrestrial conditions on human factors. (Vanderdonckt

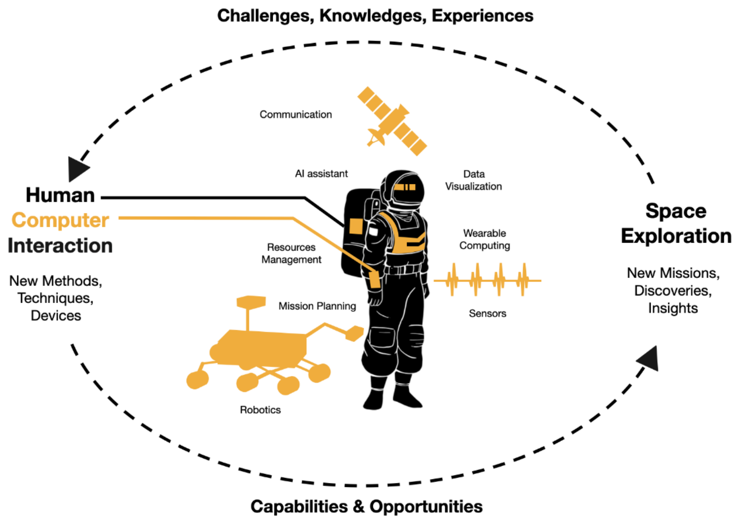


Figure 3.1: Areas of Human-Computer Interaction for Space Exploration (Pataranutaporn et al., 2022)

et al., 2024).

3.2 User Experience

The term 'user experience' (UX) can refer to a variety of things, including usability, beauty, and the hedonic, affective, and experiential aspects of using technology Hassenzahl and Tractinsky, 2006.

UX goes beyond basic usability and functionality to encompass users' emotions and perceptions before and during their interaction with a product. As UX aims to capture all aspects of user behaviour, elements such as effectiveness and efficiency are naturally included. This aligns with the approach of many practitioners, who view usability as just one component of the broader concept of UX Allam et al., n.d.

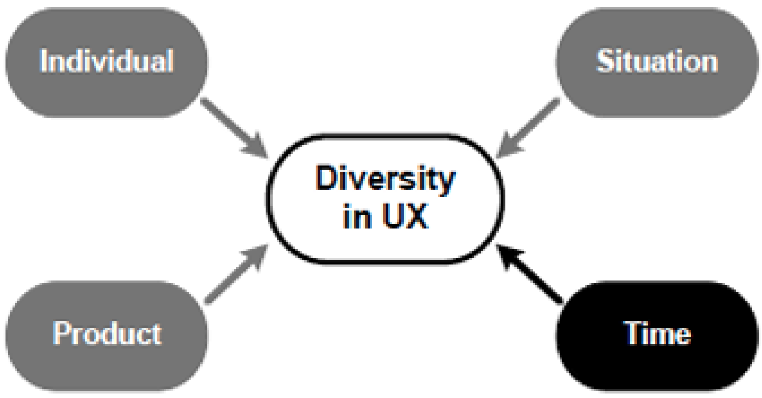


Figure 3.2: Sources of diversity in User Experience (Karapanos et al., 2008)

According to Jordan, 2000, a product has to provide useful and usable functionality before hedonic aspects, such as beauty and stimulation, can take effect. In contrast, Karapanos et al., 2008 assume that the importance of these different qualities can vary with several factors that are represented on the 3.2 and we will describe here:

- **Individual differences** will create variations in the attachment to the qualities of the product.
- The type of **product** will be a concern.
- The **Situations** in which the product is utilized are likely to influence the perception and attachment to its qualities.
- The change of experience over **Time** is an important aspect too. For instance, the more we use a product the more our perception of the qualities of the product will evolve.

For Law et al., 2009, UX is distinct from other experiences and can be defined as the Uuser experience of products, systems, services and objects that a person interacts with through a user interface. These may take the form of tools, knowledge systems, or entertainment services. Face-to-face interaction between humans is not within the remit of UX, unless there is a man-made user interface involved in the interaction. This can be showed on Figure 3.3.

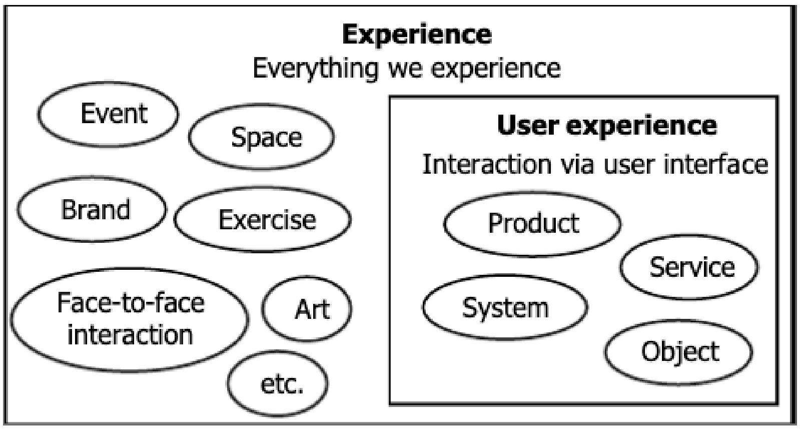


Figure 3.3: UX in relation to other experiences that we can study (Law et al., 2009)

3.3 Interaction design Framework for extraterrestrial applicability

Designing for HCI in extraterrestrial environments presents challenges that require new approaches. In this study, we propose describing three types of interaction design that could serve as models for designing interactive systems that extend beyond our usual terrestrial context..

3.3.1 Ability-Based Design

Ability-Based Design (ABD) is a user-centric approach that focuses on the user's capabilities rather than their limitations. Consequently, ABD is highly useful in scenarios where human capabilities are restricted, such as in space or on an exoplanet. The seven core principles of ABD design are ability, accountability, availability, adaptability, transparency, performance and context (Vanderdonckt et al., 2024).

3.3.2 Reality-Based Interaction

Reality-based interaction (RBI) is an approach that focuses on the relationship between human-machine interaction and the real world, in terms of both physical aspects and subjective perception of the environment. This study aims to explore the potential of HCI to approximate non-digital real-world interaction. Consequently, the author argues for the development of natural interfaces based on the conventions of everyday practices. This vision is based on four fundamental concepts: first, naive physics; second, bodily awareness and skills; third, environmental awareness and skills; and fourth, social awareness and skills Vanderdonckt et al., 2024.

3.3.3 Sensorimotor Realities

Sensorimotor realities (SRs) have the capacity to transform the HCI paradigm, transitioning it from a model based on information exchange to one characterised by sensory and motor experiences facilitated by technological interventions. This is particularly relevant in the context of extreme environments, where technological mediation is ubiquitous and encompasses all aspects of perception and action (Vanderdonckt et al., 2024).

3.4 M.A.R.S. UClouvain

In the years to come, space exploration is set to become a race for many countries and companies. For example, both SpaceX and Blue Origin have succeeded in sending civilian crews into space (Pataranutaporn et al., 2022). This illustrates that space exploration will not be limited to astronauts, but will be open to the general public Pataranutaporn et al., 2022.

However, we should ask ourselves what the next main step in space exploration should be. Mars is obviously at the forefront of everyone's mind due to its proximity to and similarity with Earth. This presents opportunities for research and comparative studies (Vanderdonckt et al., 2024).

In the context of HCI, creating new types of interactive systems on Mars presents an innovative challenge that could contribute to space exploration in general. Although it is impossible to conduct evaluations on Mars itself, we can attempt to recreate its physical environment on Earth in order to continue testing and experimentation. The Mars Desert Research Station (MDRS) in Utah is the best example of this, as its geologic landscape is comparable to that of Mars (Vanderdonckt et al., 2024).

Every year, UCLouvain sends a team of students and researchers to conduct scientific experiments and carry out geological fieldwork in preparation for future crewed missions to Mars. M.A.R.S. is a non-profit organisation whose name is an acronym that describes the project's main goal: Mars Analog Research Simulation ("M.A.R.S. UCLouvain - Martian simulation", n.d.).

Chapter 4

Materials and methods

4.1 Evaluation of User Experience

Several methods can be used to evaluate the user experience (UX) of a product. The most widely used approaches within this field are usability tests, expert reviews, and UX questionnaires (Schrepp and Thomaschewski, n.d.). This study is based on the questionnaire method. Indeed, the use of questionnaires facilitates the collection of data from a greater number of users and the subsequent conversion of their subjective impressions into numerical scale values. These scale values are indicative of the perceived quality of the product's UX within the target demographic (Schrepp and Thomaschewski, n.d.).

The User Experience Questionnaire (UEQ) is a validated and widely used UX questionnaire. It contains 6 UX scales: Attractiveness, Efficiency, Perspicuity, Dependability, Stimulation and Novelty. Therefore, the UEQ is a semantic differential that utilises a 7-point Likert scale for the responses (Schrepp and Thomaschewski, n.d.).

However, the present study is primarily based on the UEQ+ questionnaire, which contains a greater number of scales and thus provides enhanced coverage of UX aspects. The following 12 scales are employed, accompanied by concise descriptions of each (Schrepp and Thomaschewski, n.d.; Vanderdonckt et al., 2024):

- **Attractiveness** measures the overall impression from the product. Do users like or dislike the product?
- **Efficiency** measures the users' impression that they can complete their tasks without unnecessary effort.
- **Perspicuity** measures user's impression about how easy it is to become familiar with the product. It also includes the process of learning to use it.
- **Dependability** measures the user's impression that they are in control of the in-

teraction.

- **Stimulation** evaluates how exciting and motivating the product feels during its utilization.
- **Novelty** measures the degree to which the product feels innovative and creative.
- **Trust** evaluates user's impression that data entered into the product is safely processed and cannot be used to the detriment of the person providing it.
- **Adaptability** evaluates the user's impression about how easily the product can be adapted to specific preferences.
- **Usefulness** measures the user's perception of the degree to which the product brings advantages to the user.
- **Visual aesthetics** evaluates the degree to which the product has a nice and appealing visual look.
- **Intuitive use** measures user's impression about the capability of the product to be used immediately without training, instructions, or help from others.
- **Trustworthiness** of content measures the impression of participants about the information provided by the application being reliable and accurate
- **Haptics** represents the feeling resulting from touching and physically manipulating product.
- **Value** evaluates if the system's design convey a professional and high-quality appearance.

To have a good comparison the scales were employed in Earth conditions before the mission (day 0) and in Mars conditions at different time points represented by four (Sol 4), eight (Sol 8), and twelve Martian days (Sol 12). Subsequently, Mars' experience was aggregated and compared to Earth into one measure of UX transferability adapted to the specific application contexts involving UX measure α and planet π, as follows:

$$\mathrm{UXT}(\alpha, \pi) = \frac{\text{Measure } \alpha \text{ evaluated on planet } \pi}{\text{Measure } \alpha \text{ evaluated on Earth}} \tag{4.1}$$

Utilising the aforementioned formula, UXT (α = Efficiency, π = Mars) signifies, for instance, the relative perceived efficiency of interactions conducted with a particular computer system on Mars, relative to the corresponding user experience on Earth.

4.2 Internal consistency of the study via Cronbach's Alpha

Cronbach's alpha is an internal consistency coefficient commonly used in psychometrics and research methodology. It denotes the extent to which a set of items or questions in a scale or test consistently measures the underlying construct (Griffin, 2024).

It ranges between 0 and 1, the higher the values of this coefficient, the better will be the consistency inside the test or the study. It can be calculated via this formula (Griffin, 2024):

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k}\sigma_{Y_i}^2}{\sigma_X^2}\right) \tag{4.2}$$

where:

- α : Cronbach's Alpha
- k : Number of items in your scale
- $\sum \sigma_{Y_i}^2$: Summated variance of each item
- σ_X^2 : Variance of total scores

In other words:

- Zero indicates an absence of correlation between the items. The response to one question does not provide information about a response to another question.
- One indicates a perfect correlation between the items. The value of one response provides complete information about the other items (Frost, 2022).

The benchmark values for consistency are set to 0.7. At this level and above, the items demonstrate sufficient consistency to indicate the measure is reliable (Frost, 2022). However, an elevated Alpha value may occasionally be indicative of a redundant series of questions (Frost, 2022).

Chapter 5

Results analysis

5.1 Romie

Last year, crew members evaluated the user experience of `Romie` (Saint-Guillain et al., 2023), a management application designed for advanced modelling and scheduling. It was selected for its ability to optimise astronaut scheduling during missions and serves as a critical component of operations in hostile environments (Vanderdonckt et al., 2024). This empirical study was conducted within the broader framework of evaluating interaction design for extraterrestrial contexts.

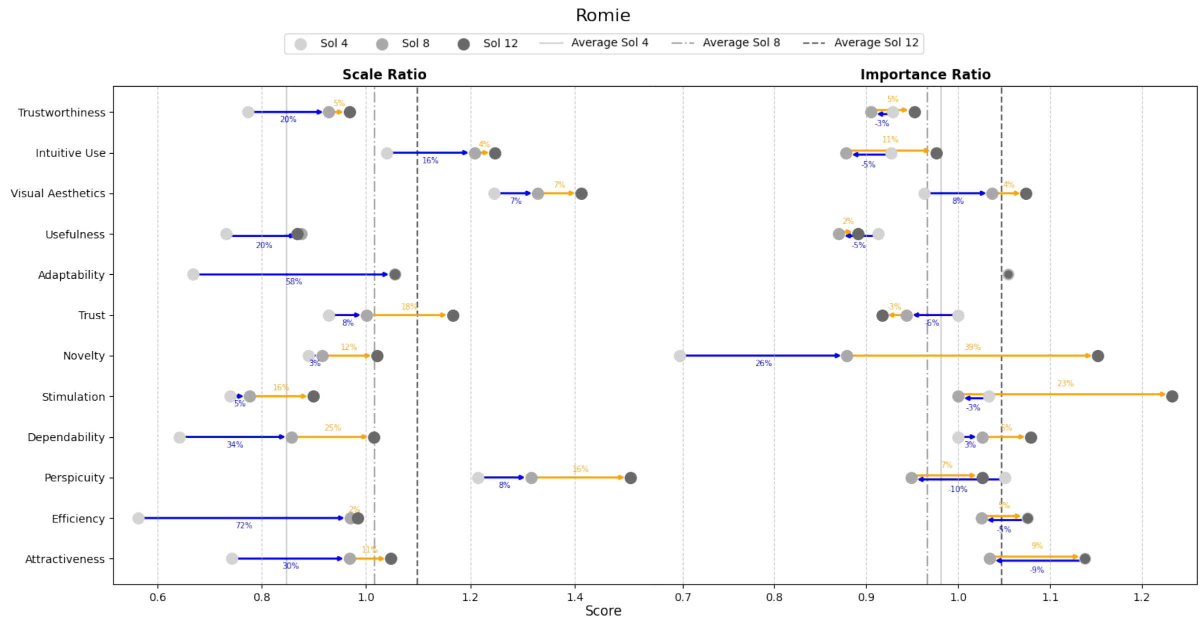
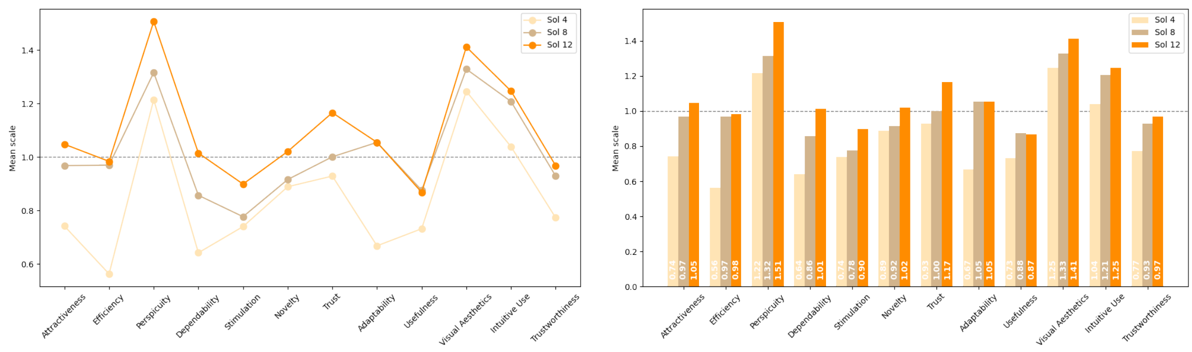


Figure 5.1: Romie: dumbbell plot of the evolution of user experience (scale) and importance ratios after four, eight, and twelve days.

Figure 5.1 illustrates the evolution of multiple *UX Transferability (UXT)* ratios and their perceived importance over four, eight, and twelve Martian days (*Sols*). UXT ratios express relative user experience on Mars compared to Earth (*Sol 0*, which is normalised to 1.0), enabling the comparison of interaction quality across environmental contexts.

5.1.1 UX Performance evolution

The evolution of UXT performance ratios reveals three adaptation phases: an initial decline at *Sol 4*, a substantial recovery by *Sol 8*, and marginal gains by *Sol 12*. This pattern illustrates that while user adaptation plays a role, certain UX dimensions are more influenced by environmental mediation than by personal adjustment.



(a) Temporal evolution (b) Performance score comparison

Figure 5.2: Romie: evolution of UX performance over the mission

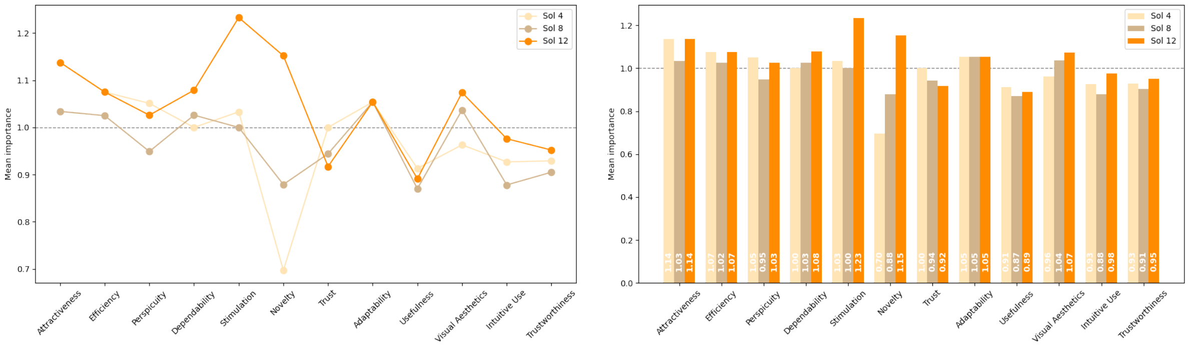
As shown in Figure 5.2, *Efficiency* exhibited the most persistent underperformance, falling to 0.56 on *Sol 4* and reaching only 0.98 by mission end. This sustained 44% deficit suggests that efficiency might be constrained by fundamental environmental factors such as limited mobility in suits. Such constraints hinder the interface's ability to support quick task completion, pointing to a need for systemic design changes rather than user training alone.

In contrast, *Perspicuity* began strong and continued to improve (from 1.21 to 1.51), suggesting that the Martian operational context, despite its complexity, supports enhanced comprehension and learning. Likewise, the rise in *Visual Aesthetics* scores implies sustained visual engagement and motivation, even in psychologically demanding settings.

Similarly, *Adaptability* exhibited the most dramatic recovery pattern, increasing by 58% between *Sol 4* and *Sol 8*. This rapid improvement suggests that users can develop effective workaround strategies for environmental limitations. However, initial constraints severely impact the system's customisation capabilities.

5.1.2 Perceived importance evolution

Perceived importance ratios, as illustrated by Figure 5.3, reveals significant changes is in user priorities over time. Several attributes — including *Trust*, *Usefulness*, *Intuitive Use*, and *Trustworthiness* — were consistently rated below their Earth baseline. These findings may reflect a shift in priorities in hostile environments: when basic functionality and operational survival dominate, conventional trust indicators and long-term value perceptions become less central.



(a) Temporal evolution

(b) Perceived importance comparison

Figure 5.3: Romie: evolution of perceived importance dynamics over the mission

Notably, *Novelty* increased in importance from 0.7 to 1.15 (+64%), with the steepest gain between *Sol 8* and *Sol 12*. This rise suggests a growing need for psychological stimulation, suggesting that prolonged isolation increases reliance on engaging or novel interface features. *Stimulation* reached the highest importance value (1.23), reinforcing this trend.

Additionally, the values for *Attractiveness*, *Efficiency*, *Dependability* and *Stimulation* all demonstrated a consistent increase that exceeded the baseline value of Sol 0. This could suggest a progressive realignment of user expectations with environmental demands.

5.1.3 Strength and limitations

Given the results observed, it is apparent that Romie is effective in hostile environments, sustaining visual engagement. The superior *Perspicuity* performance indicates that the interface design successfully accommodates the characteristics of a focused attention environment in highly challenging operational situations. Furthermore, the strong *Visual Aesthetics* ratings suggest that user motivation is effectively maintained throughout extended mission durations.

However, as illustrated in Figure 5.2, the persistent deficits in Efficiency and Stimulation represent a critical weakness. The performance analysis demonstrates that, whilst the majority of dimensions achieve or approach Earth parity by *Sol 12*, *Efficiency*, *Stimulation*, *Usefulness* and *Trustworthiness* remain below *Sol 0* performance. Despite the fact that *Usefulness* and *Trustworthiness* are not perceived as important as *Efficiency* and *Stimulation*, as seen in Figure 5.3, the performance gap that remains throughout the mission may indicate fundamental interaction challenges that user adaptation cannot resolve.

We can also observe a diminished significance of conventional trust indicators may signify an incompatibility with the hostile environment mindset. Moreover, the mounting importance of *Novelty* and *Stimulation* suggests that static interface designs may become increasingly unfit for purpose during extended missions.

5.1.4 Design adaptation strategies

The temporal patterns exhibited in Figure 5.2a and 5.3a suggest that efforts to enhance these systems should prioritise compensating for persistent environmental constraints rather than relying on user adaptation.

Additionally, the system should incorporate dynamic interface elements that evolve throughout the mission to address the growing importance of psychological engagement factors demonstrated in Figure 5.1. This could encompass progressive feature revelation or adaptive visual themes that respond to the increasing value of *Novelty* and *Stimulation* over time.

Finally, the performance analysis represented by Figure 5.2 indicates that pre-mission training should emphasise efficiency workarounds rather than expecting full performance recovery, while leveraging the system's demonstrated strength in supporting complex learning under Mars conditions. This approach aligns with the principles of Ability-Based Design, which emphasise adapting system behaviour to users' capabilities rather than expecting users to overcome environmental constraints (Wobbrock et al., 2011). In contexts where mobility and iput are limited, Romie's interface to amplify users' existing abilities.

5.2 EZExFit

`EZExFix` is a low-cost, rapidly deployable and user-friendly external fixator designed for the treatment of tibial shaft fractures. Engineered for extreme environments, it is particularly relevant for extraterrestrial missions, where medical interventions may need to be conducted under physical constraints (such as wearing suits, experiencing low gravity or being in isolation) and under psychological pressure.

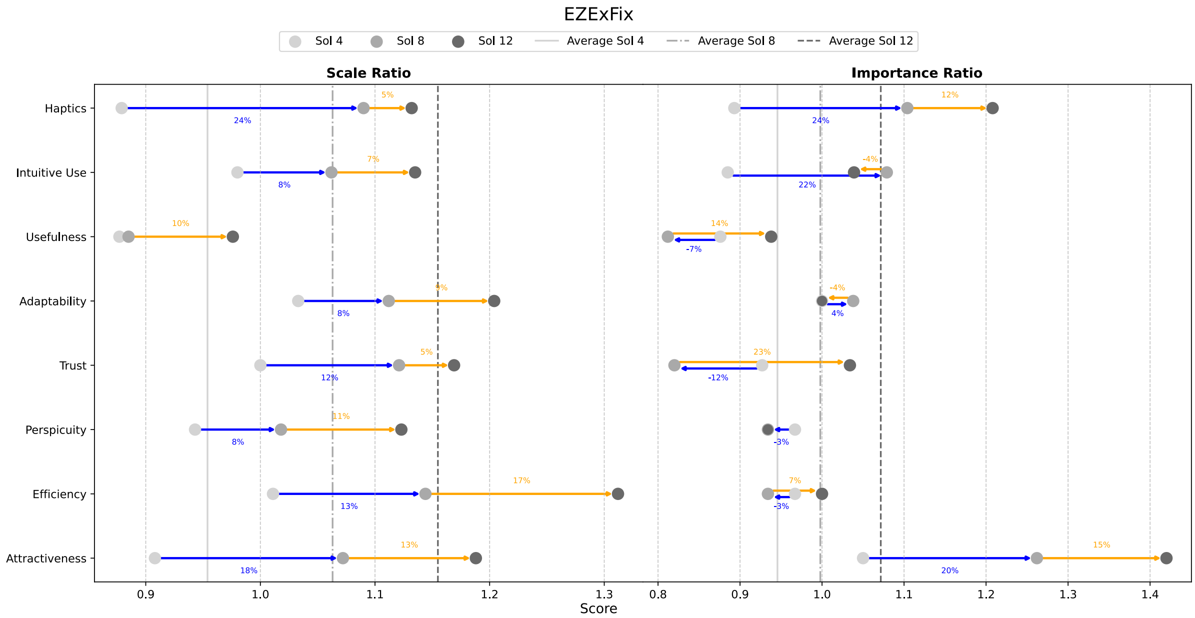
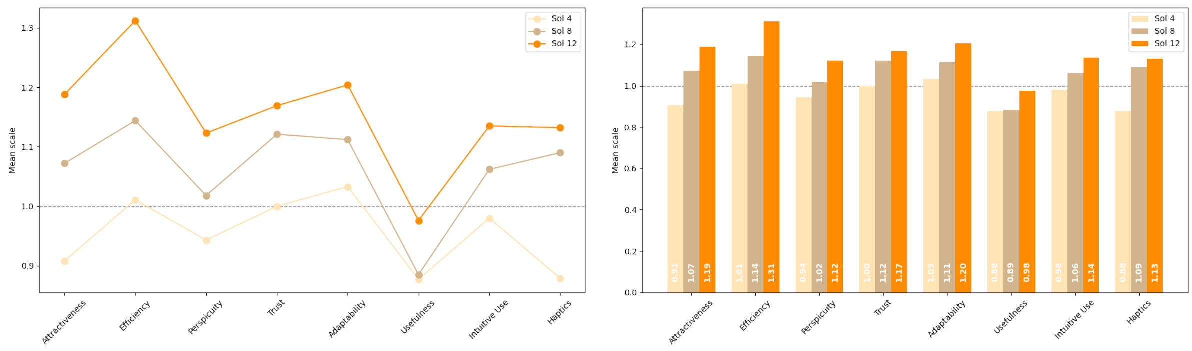


Figure 5.4: EZExFix: dumbbell plot of the evolution of user experience (scale) and importance ratios after four, eight, and twelve days.

The methodology employed in the evaluation process was consistent with the approach adopted in the evaluation of `Romie`. The study was conducted through a multi-session UX study during a Mars analogue mission. The study focused on how astronauts adapted to and perceived the device in various conditions. Figure 5.4 provides a visual representation of the resulting *UXT* ratios and their perceived importance over four, eight and twelve Martian days.

5.2.1 UX Performance evolution

The temporal evolution of UX performance ratios, as demonstrated in Figure 5.5a for `EZExFix`, exhibits a pronounced learning curve in the initial stages, subsequently followed by continuous and consistent improvement throughout the mission. The initial phase (*Sol 4*) mirrors the initial operational difficulty, which is presumably attributable to the novelty of the medical interface and the elevated stress context in which it was introduced. However, in contrast to `Romie`, `EZExFix` revealed a more linear progression in terms of UX performance. This is particularly evident in the case of *Efficiency*, which began at 1.01



(a) Temporal evolution (b) Performance score comparison

Figure 5.5: EZExFix: evolution of UX performance over the mission

on *Sol 4* and increased to 1.31 (+29%) by *Sol 12*. This steady growth indicates that the system's design, was well-aligned with user expectations and became increasingly intuitive even under constrained conditions.

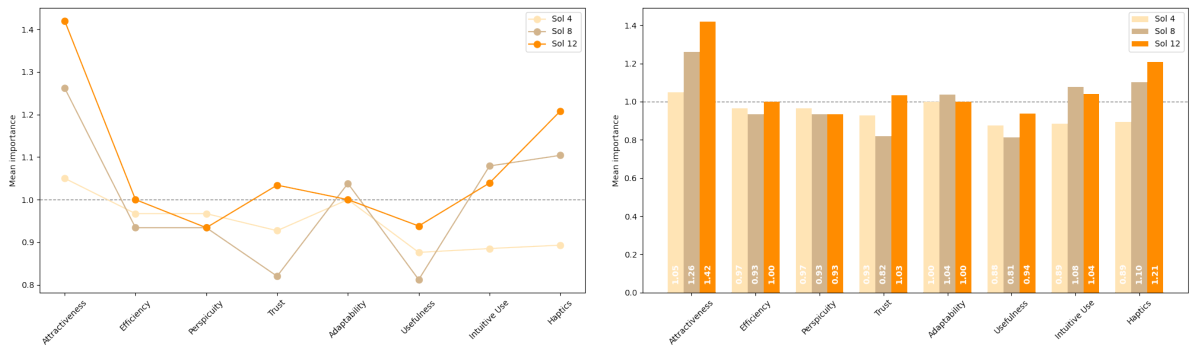
A similar trend was observed in *Perspicuity*, which improved from 1.02 to 1.12, indicating an enhancement in the ease with which users could navigate and utilise the system. Furthermore, *Haptics* have been shown to improve from 0.88 to 1.13, which indicates effective accommodation of physical components despite the constraints imposed by equipment and limited mobility.

Trust increased from 1.00 on *Sol 4* to 1.17 on *Sol 12*, suggesting that repeated exposure to operational stress reinforces perceptions of reliability and stability. It is notable that *Attractiveness* increased from 0.91 to 1.19, implying an increased appeal of the design of the system over time. This tendency is consistent with the notion that familiarity tends to induce both comfort and positive sentiment, particularly once initial functional challenges are surmounted.

The only attribute that remained below the *Sol 0* reference level was *Usefulness*, which peaked at 0.98 by *Sol 12*. This suggests that doubts about the system's role or effectiveness were present throughout the mission.

5.2.2 Perceived importance evolution

While UX performance rose consistently, importance ratios tell a more nuanced story of shifting priorities. Haptics, initially rated at 0.89 on Sol 4, rose to 1.21 by Sol 12, highlighting the growing recognition of tactile feedback as a critical aspect of system performance in physical environments. Similarly, Attractiveness exhibited the sharpest increase in perceived importance (from 1.05 to 1.42), indicating that psychological and affective dimensions became increasingly relevant as the mission progressed.



(a) Temporal evolution (b) Perceived importance comparison

Figure 5.6: EZExFix: evolution of perceived importance dynamics over the mission

As depicted in Figure 5.6, the perceived importance manifested more nuanced shifts in priorities. The two attributes that demonstrated the most notable increase in comparison to the others were *Haptics* and *Attractiveness*. The former was initially rated at 0.89 on *Sol 4* and rose to 1.21 by *Sol 12* (+36%), while the latter exhibited the most marked increase in perceived importance, rising from 1.05 on *Sol 4* to 1.42 on *Sol 12* (+35%). This highlights the growing recognition of haptic feedback as a crucial aspect of system efficiency in hostile environments, and that psychological and emotional dimensions become more significant as the mission progress.

Efficiency, *Perspicuity*, and *Adaptability*, although relatively stable, remained around *Sol 0* baseline (varying between 0.93 and 1.04), which reflects consistent expectations across contexts. Initially, Usefulness saw a decline in importance (0.88 at *Sol 4* to 0.81 at *Sol 8*), but recovered to 0.94 by *Sol 12*. This progression suggests that as the system proved effective, users reconsidered its operability.

The lowest and most variable score was observed for *Trust* at 0.82 on *Sol 8*, indicating a temporary decline in confidence, but the recovery to 1.03 by *Sol 12* supports the system's capacity to gain user trust over time. *Intuitive Use* showed a slight decrease after a peak at 1.08 on *Sol 8*, reaching 1.04, which may be attributable to a need for greater control and comprehension as the mission progresses.

5.2.3 Strength and limitations

As demonstrated in the above sections, the `EZExFix` system showed a robust UX profile in physical and hostile environments. Its key strengths are consistent and substantial improvement in *Attractiveness*, *Efficiency* and *Haptics*. The performance trajectory of *Perspicuity* confirms that the system supports rapid learning and effective usage, even under mission stress. Moreover, the increasing appeal and perceived significance of the visual and tactile dimensions indicate that the system proposes practical functionality as well as a degree of psychological reassurance and engagement.

Nevertheless, certain limitations did become apparent. *Usefulness* of the scale consistently lagged behind other attributes in both scale and perceived importance. This finding suggests that the operational role may not have been as evident or acknowledged in the initial phases of the mission. In a similar way, the initial perceptions of *Haptics* and *Intuitive Use* were found to be somewhat limited. This finding suggests the presence of initial challenges in manipulating the device while wearing equipment.

5.2.4 Design adaptation strategies

The progressive elevation in both perceived performance and importance suggests that future design strategies should reinforce this positive adaptation curve via such means as modular support tools or in-context usage guide. Enhancing the emphasis on utility during the first interactions could help to reduce the initial underestimation of *Usefulness*, *Haptics*. This emphasis on haptic feedback and the ergonomic grip is consistent with the principles of Sensorimotor Realities, which emphasise the significance of mediated perception and motor function in confined environments (Vatavu, 2022). By enhancing tactile affordances, the system effectively compensates for diminished direct physical control, such as that caused by gloved manipulation or limited dexterity.

When considering the rise in importance of aesthetic and psychological factors, such as *Attractiveness* and *Intuitive Use*, future designs could take into account the incorporation of adaptive features that respond to mission duration and user familiarity. These enhancements would ensure the maintenance of high functionality, and sustain cognitive engagement and trust in the system throughout the duration of the mission.

Finally, it is important to emphasise that the system's usability under mobility constraints should continue to guide interface decisions. Furthermore, the reinforcement of positive tactile feedback and the assurance of grip stability during glove use can ensure that the strong *Haptics* scores observed at the conclusion of the mission become a baseline for the system.

5.3 Tap Strap 2

Tap Strap 2 is a single handed all-in-one wearable device that functions as a keyboard, mouse and air gesture controller. The device was designed for use in both gesture-based and touchless input, rendering it particularly relevant for use in confined and hostile environments. In such contexts, conventional input devices may be impractical due to the equipment required to conduct the missions. The objective of the wearable is to provide intuitive and versatile control for interaction with computer systems whilst minimising the need for surface-based interfaces.

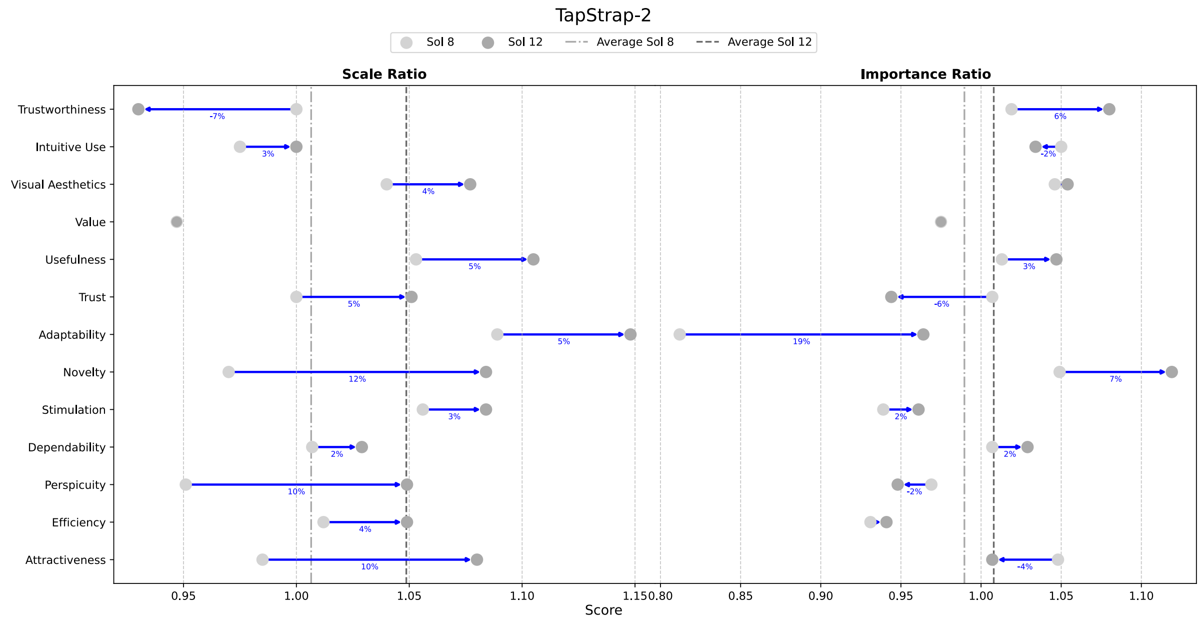
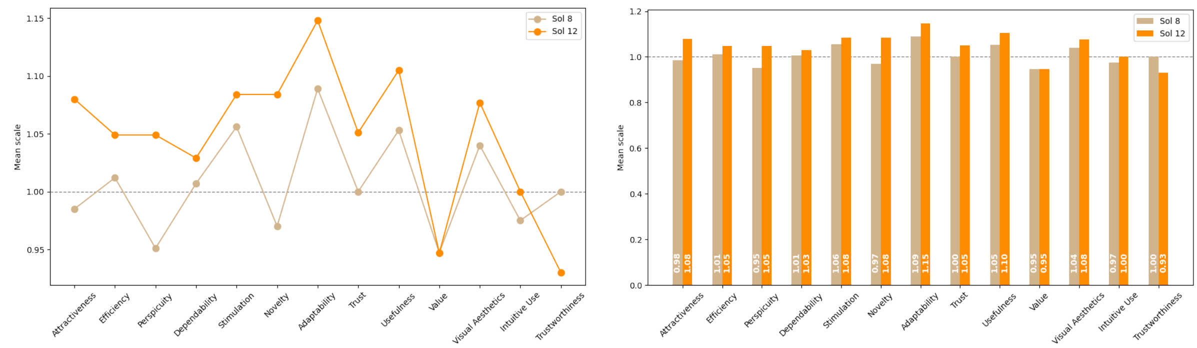


Figure 5.7: Tap Strap 2: dumbbell plot of the evolution of user experience (scale) and importance ratios after eight and twelve days.

The methodology employed in the evaluation process differs from the one employed with `Romie` and `EZExFix` systems. Instead of employing Sol 0 as the baseline for UX performance, it was established at Sol 4. Figure 5.7 provides a visual representation of the resulting UXT ratios and their perceived importance over eight and twelves day.

5.3.1 UX Performance evolution

As illustrated by Figure 5.8, the progression of UX performance for `Tap Strap 2` exhibits a consistent positive trend across the majority of the evaluated attributes. As indicated by the Sol 4 baseline, most metrics demonstrated increase performance, most notably *Adaptability*, *Novelty* and *Stimulation*. *Adaptability* improved to 1.09 at *Sol 8* up to 1.15 at *Sol 12*, suggesting the users' increasing capacity to integrate `Tap Strap 2` into their environment. Conversely, Novelty component exhibited an increase to 0.97 at *Sol 8* up



(a) Temporal evolution
(b) Performance score comparison

Figure 5.8: Tap Strap 2: evolution of UX performance over the mission

to 1.08 at *Sol 12*, which could indicate that the device's interactive nature maintained its appeal and continued to engage users as their proficiency with it increased.

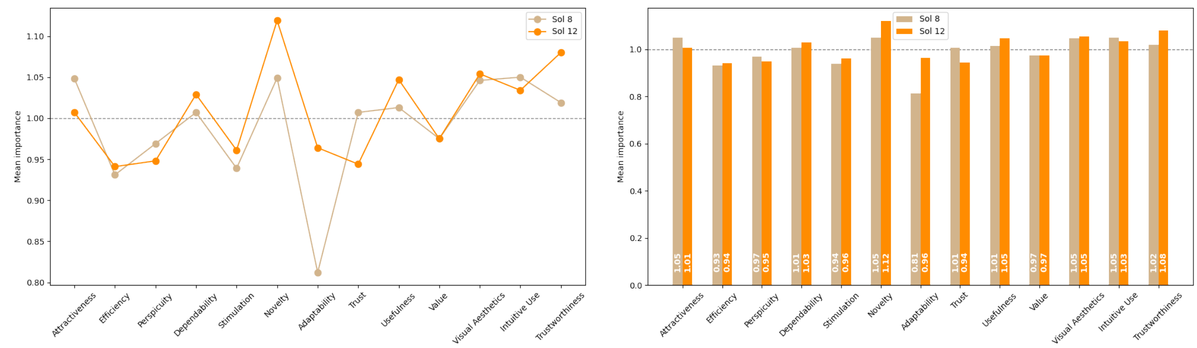
Stimulation showed steady gains, growing to 1.06 at *Sol 8* and to 1.08 at *Sol 12*, suggesting that the device contributed positively to user engagement and remained stimulating even in repetitive or high-pressure conditions. This is a significant result in the context of long-term missions, where psychological sustainability is a critical factor in the system's adoption.

Furthermore, two core functional dimensions, namely *Perspicuity* and *Efficiency*, exhibited moderate improvements. A significant enhancement in clarity was observed from *Sol 8* to *Sol 12*, with an increase from the baseline value of 0.95 at *Sol 8* to 1.05 at *Sol 12*. This finding suggests that users gradually became more proficient in comprehending and utilising the system with repeated usage. This observation is further evidenced by *Efficiency*, which exhibited an enhancement from 1.01 at *Sol 8* to 1.05 at *Sol 12*.

Nonetheless, certain attributes demonstrated either negligible or declining performance. *Trustworthiness* showed a decrease to 0.93 by Sol 12, suggesting the emergence of doubts concerning the reliability or perceived integrity of the device over time. In a similar manner, *Value* remained static and below the baseline, maintining a constant 0.95. This may indicate a diminished vision of the pratical utility of the system, as perceived by the users.

5.3.2 Perceived importance evolution

The evolution of perceived importance across the different attributes exhibited a more variable pattern than for performance, as illustrated in Figure 5.9. Attributes such as *Novelty* and *Trustworthiness* gained significant perceived importance, by increasing up to 1.12 and 1.08 by *Sol 12* respectively. This may reflect users' growing appreciation for features supporting engagement and system integrity as interaction demands evolve.



(a) Temporal evolution (b) Perceived importance comparison

Figure 5.9: TapStrap-2: evolution of perceived importance dynamics over the mission

Usefulness also experienced a modest increase up to 1.05 at *Sol 12*, which can be aligned with the performance increase observed earlier.

Conversely, several foundational dimensions either declined or stagnated. For instance, *Efficiency* decreased to 0.93 at *Sol 8* and increased a bit to 0.94 at *Sol 12*. The same phenomenon can be observed with *Perspicuity*, which decreased down to 0.95 at *Sol 12*. This observed dissociation between performance and importance may be indicative of a cognitive shift wherein users, having become accustomed to the system, have redirected their focus towards more experience-based or hedonic features.

Adaptability exhibited a distinctive trajectory, with an initial decline to 0.81 at *Sol 8* follow by a sharp increase to 0.96 at *Sol 12*. While it remains below the baseline, this curve may be indicative of a temporary discrepancy between user expectation and system flexibility, which is likely to be resolved through growing familiarity or through the implementation of workaround strategies.

Interestingly, *Stimulation* and *Intuitive Use*, two attributes that could be expected to rise in importance over time in confined or hostile environments, instead fluctuated with modest increases, reaching 0.96 and 1.03 respectively at *Sol 12*. These figures suggest that, while users appreciated the experiential contributions of `Tap Strap 2`, such aspects did not rise to critical significance compared to other functionality-focused dimensions.

5.3.3 Strength and limitations

Following the observations outlined above, `Tap Strap 2` demonstrate notable versatility and psychological engagement in constrained or hostile environments. Its performance in dimensions such as *Adaptability*, *Novelty*, and *Stimulation* reflects a string alignment with the demands of such settings. These strengths are especially valuable in extraterrestrial environments, where it is crucial to minimise cognitive fatigue and maintain user morale.

The steady increase in *Efficiency*, *Perspicuity*, and *Usefulness* further supports the system's viability as a reliable input mechanism for such missions. It can be noted that it's wearable design does not appear to impede learning or task execution, even in the presence of persistent environmental constraints.

Nevertheless, the persistent deficiency in perceived *Value* and the decline in *Trustworthiness* highlight potential constraints in long-term reliability and most importantly user confidence. These issues may be attributable to factors such as gesture recognition inconsistencies, calibration drifts or a lack of clear feedback during its usage. Furthermore, while *Novelty* was consistently appreciated, its limited translation into perceived utility suggests a discrepancy between engagement and practical contribution to workflow.

5.3.4 Design adaptation strategies

In consideration of the performance and perceived importance evolution observed, `Tap Strap 2` could benefit from design enhancements focused on transparency, feedback, and reliability over time. The utilisation of real-time feedback mechanisms, or usage confirmation signals, could prove effective in enhancing perceived reliability and mitigating user uncertainty.

To enhance the perceived *Value* of the system, designers could incorporate adaptive usage metrics or productivity indicators that emphasise the system's practical benefits. These measures would facilitate the integration of the experiential and operational aspects of the device.

Furthermore, given the significance of psychological stimulation and novelty, a potential enhancement of `Tap Strap 2` could involve incorporating responsive visual elements that adapt based on usage context or mission phase. This approach would facilitate the sustained engagement and interest of the user, whilst simultaneously enhancing the adaptability of the system. This approach also aligns with the adaptability principle of Ability-Based Design, as it enables users to customise inputs to suit their individual dexterity and the complexity of the task at hand over time Wobbrock et al., 2011.

Overall, the wearable nature and versatility of `Tap Strap 2` position it as a highly promising tool for constrained operational environments.It is imperative to acknowledge that the continuous refinement of trust and value-oriented features is essential to achieving the full potential of this system in constrained or hostile environments.

5.4 Parrot drone

Utilising a Parrot drone for UX analysis introduces a substantial shift when compared to the previous systems that have been examined within the scope of this thesis. Indeed, the device is a physical robotic system designed for remote manipulation rather than being a wearable or software-driven system. The evaluation process focused on the execution of controlled flight tasks across various mission stages. The drone's interface demands a combination of spatial reasoning, coordination, and responsiveness, all whilst operating within the constraints of different equipment and an overall hostile environment.

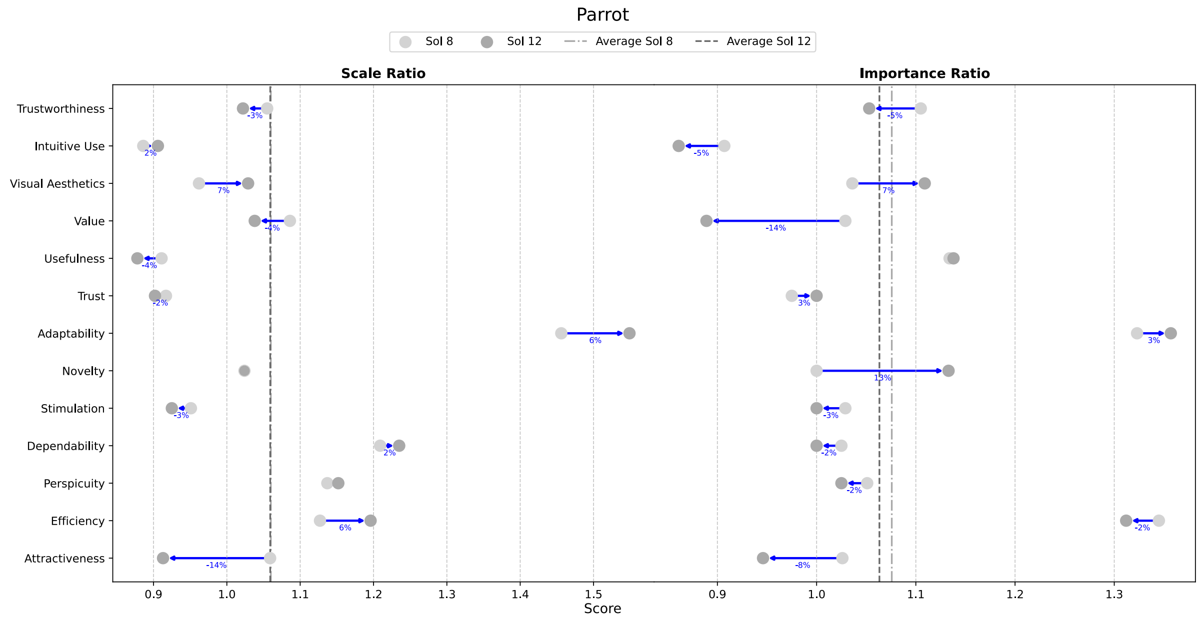
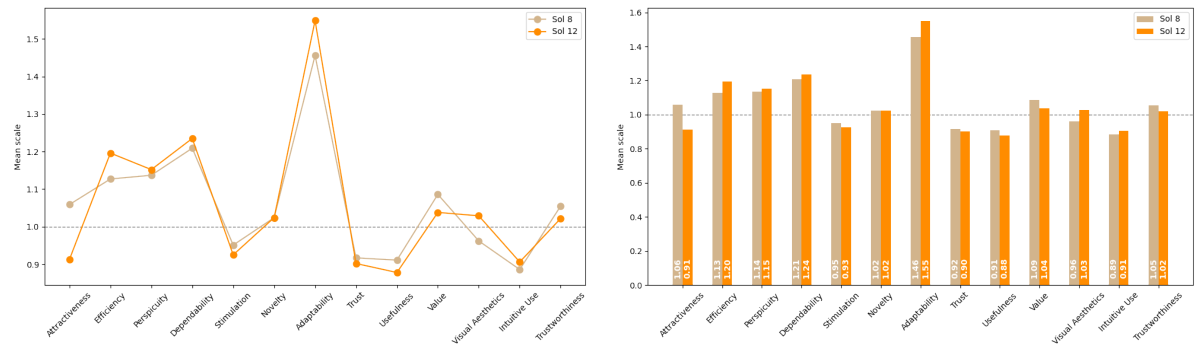


Figure 5.10: Parrot drone: dumbbell plot of the evolution of user experience (scale) and importance ratios after eight and twelve days.

As with `Tap Strap 2`, the analysis takes Sol 4 as the baseline. Figure 5.10 depicts the evolution of user experience transferability (UXT) across performance and perceived importance metrics. The following sections explore the observed trends and their implications over time.

5.4.1 UX Performance evolution

The Parrot drone demonstrated a more polarised evolution of performance over time. As shown in figure 5.11, specific attributes such as *Efficiency*, *Perspicuity* and *Dependability* exhibited significant increases. In fact, *Efficiency*, showed a steep increase from its baseline to 1.13 at Sol 8 and up to 1.2 at *Sol 12*. This progression is indicative of a sharp but manageable learning curve, where users progressively improve their proficiency in execution of precise manoeuvres and control sequence.



(a) Temporal evolution

(b) Performance score comparison

Figure 5.11: Parrot drone: evolution of UX performance over the mission

As enounced before, *Perspicuity* and *Dependability* followed similar trends, reaching 1.15 and 1.24 respectively at *Sol 12*. These results indicate an improvement in both the comprehensibility of the system and the perception of its stability, despite the initial unfamiliarity of the user, and the physical disconnection between the user input and the drone action.

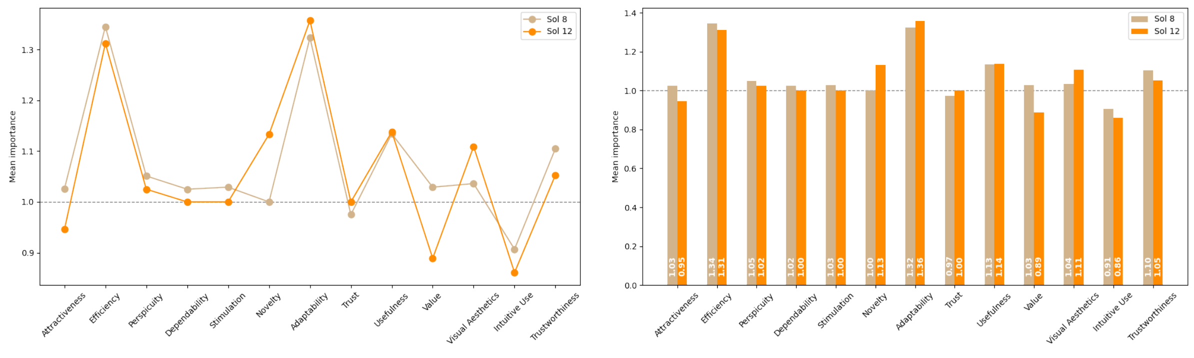
On the other hand, several dimensions exhibited stagnation ordecline. Notably, Usefulness and Trust decreased down to 0.88 and 0.90 respectively. This could suggest that, while the drone was technically operational, it lacked a consistent sense of operational value across the various phases of the mission. *Similarly*, *Trustworthiness*, despite an initial increase, settled at a modest 1.02 by *Sol 12*. This relatively flat result suggests that users may have had limited confidence in the overall system.

Stimulation and *Attractiveness* also showed a slight decline over time. Both dipped below the baseline, with *Stimulation* decreasing to 0.93 and *Attractiveness* to 0.91 by *Sol 12*. This pattern is distinct from that observed in other systems, such as `EZExFix` and `Tap Strap 2`, where stimulation levels increased in proportion to the user's level of proficiency.

However, *Adaptability* demonstrated the most significant improvement in comparison with the other attributes. Indeed, an increase was observed, reaching 1.46 by *Sol 8*, and up to 1.55 by *Sol 12*, thus demonstrating the capacity of the system to adapt to the mission's constrained and hostile environment.

5.4.2 Perceived importance evolution

Perceived importance patterns, as illustrated in Figure 5.12, revealed a more mixed trajectory. Efficiency and Adaptability saw the most substantial gains, increasing by over 30% at Sol 12 compared to the baseline of Sol 4. This suggest that users grew to appreciate flexible control and system responsiveness as they became more skilled in navigating



(a) Temporal evolution (b) Perceived importance comparison

Figure 5.12: Parrot drone: evolution of perceived importance dynamics over the mission

and trubleshooting the drone.

Perceived importance patterns, as illustrated in Figur 5.12, revealed a more nuanced trajectory. The most substantial gains were observed for *Efficiency* and *Adaptability*, which increased by over 30% at *Sol 12* compared to the baseline of *Sol 4*. This finding suggests that, as users' proficiency in operating and troubleshooting the drone increased, they came to value flexibility in control and system responsiveness.

Novelty and *Visual Aesthetics* also exhibited an increase in perceived importance, with the former reaching 1.13 at *Sol 12* and the latter reaching 1.11 at the same point. These increases may be indicative of a late-phase recognition of the drone's unique experiential value, potentially influenced by variations in tasks or the distinctiveness of flying in comparison to other forms of interaction.

Conversely, other attributes exhibited a decline. *Trust*, *Value*, and *Intuitive Use* all dropped below baseline by *Sol 12*, which is noteworthy given their significance in previous systems. It may be inferred that, although the drone system improved in operability, it did not convincingly justify its role within the larger operational context of the mission.

The perceived importance of *Intuitive Use* and *Usefulness* also declined, suggesting that despite enhanced control, the system did not offer intuitive learning or predictable responses from the users' perspective. This finding suggests a potential misalignment between the complexity of the system and the limited learning opportunities available during the mission.

5.4.3 Strength and limitations

The Parrot drone demonstrated notable strengths when it came to user skill acquisition and increasing control precision. The positive trend in both *Efficiency* and *Dependability* indicates an increase in users' proficiency in the operation of the drone. These are of paramount importance in systems that necessitate spatial and remote operations under environmental constraints.

However, it still suffered from declining perceptions of *Value* and *Usefulness*, two aspects that had remained strong in other tools studied earlier. The restricted perceived *Adaptability* and *Intuitive Use*, even at *Sol 12*, also reflect barriers of ease-of adoption, potentially stemming from the drone's external feedback loop and control complexity.

Finally, the consistent decline in *Stimulation* and *Trust* indicates that user engagement and confidence may deteriorate over time, particularly in the absence of critical utility or rewarding task outcomes. Conversely, systems such as `Tap Strap 2` have shown sustained engagement by offering unique and varied experiences. In comparison, the drone interface may have lacked the capacity to effectively integrate tasks, potentially leading to reduced user relevance.

5.4.4 Design adaptation strategies

Drawing from these observations, future refinements to the drone system should prioritise the facilitation of an intuitive onboarding process and the incorporation of more explicit, context-sensitive feedback mechanisms. The incorporation of real-time visual or haptic feedback could serve to mitigate control ambiguity, thereby promoting enhanced trust in the system's reliability. This could also help mitigate the disconnection between action and result, an issue central to Reality-Based Interaction, which stresses the need to anchor system behavior in users' real-worl expectations of physics, feedback, and control (Jacob et al., 2008).

In order to address the decline in perceived value, it is essential that the drone use case become more firmly embedded within the mission workflow. By establishing a direct correlation between the performance of drones and the operational outcomes they produce, users can more effectively assimilate their role and substantiate the allocation of attention and effort.

Furthermore, the initial rise and subsequent plateau of *Trustworthiness* suggest a need for reinforced error recovery mechanisms and performance transparency. Features such as predictive navigation, automated corrections, and instructional overlays have the potential to facilitate the learning process and mitigate confidence decline.

In summary, the Parrot drone demonstrates potential in supporting remote operations but shows deficiencies in terms of *Intuitive Use*, *Stimulation* and *Usefulness*. It occupies a niche role that requires careful design and task integration to ensure its applicability in hostile and high-stakes environments.

5.5 Harmony LLM

`Harmony LLM` is designed to assist astronauts with technical queries and operational guidance in real time. The system's natural language interface and knowledge-rich backend are designed to provide contextualised, on-demand support for mission-critical decision-making. The system was subjected to a comprehensive evaluation across multiple mission phases (*Sol 4* baseline, followed by *Sol 8* and *Sol 12*), with a particular emphasis on the evolution of its usability and perceived value within a simulated Martian environment.

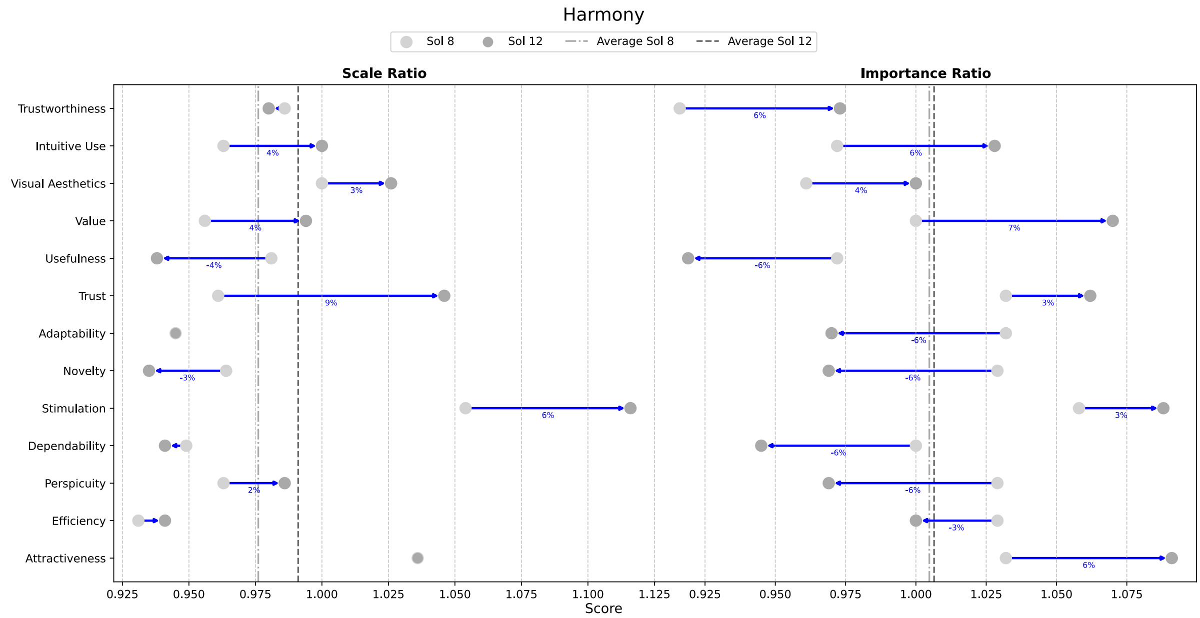


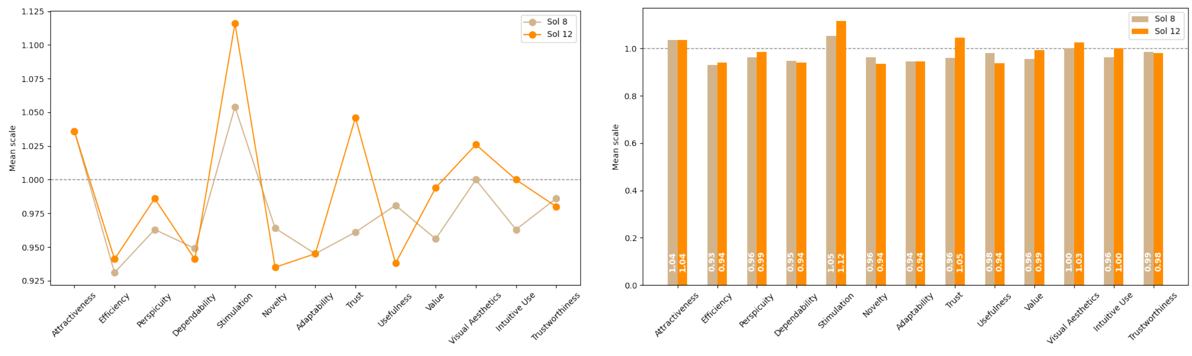
Figure 5.13: Harmony LLM: dumbbell plot of the evolution of user experience (scale) and importance ratios after eight and twelve days.

Figure 5.13 presents a consolidated view of UX scale and importance ratios. The following sections examine these results in detail.

5.5.1 UX Performance evolution

As demonstrated by Figure 5.14, `Harmony` exhibits a relatively stable and balanced UX performance. The highest score was observed in *Stimulation*, which increased to 1.05 at *Sol 04* up to 1.12 at *Sol 12*. This finding indicated that users found the system increasingly engaging over the mission, which is crucial in long duration missions where cognitive stimulation is essential.

Other positive trends can be observed in *Trust*, which improved from below the baseline at 0.96 at *Sol 8* to 1.05 at *Sol 12*. *Visual Aesthetics* and *Intuitive Use* also saw a gradual improvement, suggesting a growing level of familiarity and appreciation for the system's presentation and interface design.



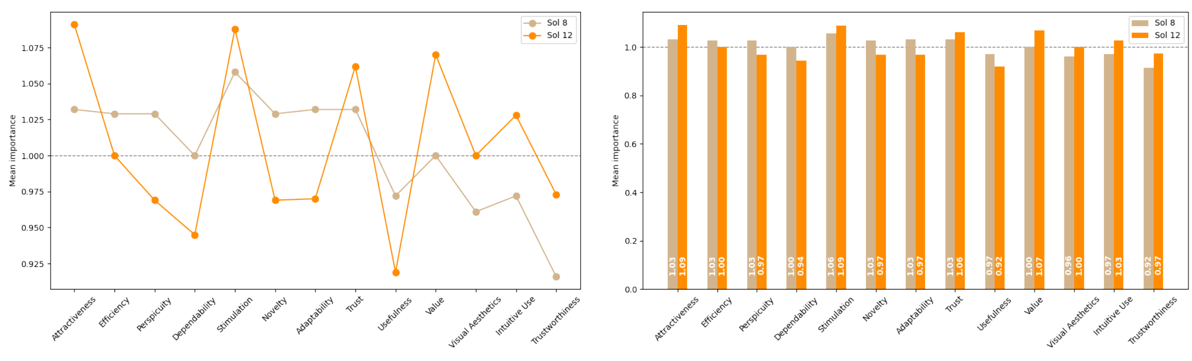
(a) Temporal evolution (b) Performance score comparison

Figure 5.14: Harmony LLN: evolution of UX performance over the mission

Generally, the performance ratios exhibited stability across the majority of the attributes, with only minor fluctuations. Namely, *Efficiency*, *Dependability* and *Adaptability* declined to 0.94 *Sol 12* respectively. These drops may indicate either interface limitations or perceived variability in system responsiveness over time.

An interesting observation is the decline in *Usefulness*, which decreased to 0.98 at *Sol 8* down to 0.94 at *Sol 12*. This could suggest that, despite the overall positive ratings assigned to each attribute, there is a possibility that the relevance of `Harmony` has diminished in users' perceptions during the mission.

5.5.2 Perceived importance evolution



(a) Temporal evolution (b) Perceived importance comparison

Figure 5.15: Harmony LLN: evolution of perceived importance dynamics over the mission

As can be observed in Figure 5.15, the perceived importance generally demonstrated a strong correlation with performance metric. However, it showed slightly more divergence than previous systems. *Stimulation* and *Trust* both rose in perceived importance up to 1.09 and 1.09 by *Sol 12*, respectively, mirroring their corresponding scale increase.

Value experienced a modest but clear increase to 1.07 by *Sol 12*, suggesting that over time, users began to recognise `Harmony`'s potential long-term utility more clearly, despite some decline in its perceived immediate usefulness.

In contrast, *Usefulness*, *Dependability*, and *Novelty* all exhibited a decline in perceived importance. This may be indicative of a transition in user focus from utility to experiential qualities as familiarity with the system increases. The decline in *Efficiency* to the baseline at *Sol 12*, despite its initial prominence at 1.03 at *Sol 8*, mirrors the trend observed in `Tap Strap 2` and `Parrot`, where experiential dimensions gained ground over time.

5.5.3 Strength and limitations

A significant strength of `Harmony` is its capacity to sustain elevated levels of engagement and trust over an extended period. The consistent rise in the concepts of *Stimulation*, *Trust* and *Value* emphasises the potential for the system to act as a psychologically supportive and reliable mechanism in hostile environments. These trends are consistent with those previously observed in `EZExFix`, where sustained user motivation and attention were equally valued.

Furthermore, the stability of *Visual Aesthetics*, along with the modest gains in *Intuitive Use*, indicates that `Harmony`'s design promotes sustained usability without inducing fatigue or complexity issues. This is a significant asset for an AI assistant that is employed across various operational contexts.

However, it is important to note several key limitations. A decline in performance and perceived importance of *Usefulness* of the system has been observed, raising concerns about the scope or integration of its capabilities in real operational workflows. In a similar manner, the plateau or decrease in *Adaptability* and *Dependability* suggests that, while the system is competent, it may not scale fluidly with more complex or evolving tasks. The limitations described are similar to those previously observed with the `Parrot` drone, where trust was partially maintained but did not fully manifest as perceived reliability.

5.5.4 Design adaptation strategies

The development of `Harmony` for higher long-term impact could be improved by integrating it more deeply into task-critical procedures. This could include the use of contextual prompts, embedded procedural guides, or dynamic responses that are tailored to the mission phase. This could potentially counteract the observed decline in both *Usefulness* and *Adaptability*.

In order to maintain and strengthen perceived *Trustworthiness*, the system could include transparent explanation features. These could take the form of confidence scores or rationale trees, which would aim to reinforce users' sense of control and understanding.

When considering the positive trend in both *Stimulation* and *Trust*, adaptive narrative framing or conversational variation could be introduced to maintain user engagement. This approach would preserve novelty and mitigate habituation.

Overall, `Harmony` is shown to be a psychologically engaging and trustworthy assistant that supports user motivation effectively. Its limitations, while not insurmountable, do highlight the necessity for enhanced procedural integration and greater adaptability to align perceived and actual utility over extended mission durations.

Chapter 6

Conclusion and perspectives

This thesis sought to explore and analyse the human-computer interaction characteristics of several interactive systems in the context of analogue Martian missions. To this end, empirical data collected at the Mars Desert Research Station (MDRS) as part of the M.A.R.S. UCLouvain initiative was analysed using the *UEQ+* framework. The assessment involved five distinct systems: `Romie`, `EZExFix`, `Tap Strap 2`, a `Parrot` drone, and `Harmony LLM`. Each system was evaluated across different mission phases.

Each system was examined through a consistent analytical process, focusing on a set of dimensions of user experience. These dimensions included core usability attributes (*Efficiency*, *Perspicuity*), experiential qualities (*Stimulation*, *Novelty*), and perception-based factors (*Trust*, *Usefulness*, *Value*). In order to facilitate cross-system comparison of the evolution of these attributes under hostile environmental conditions, the study incorporated normalized transferability ratios based on baseline performance (either *Sol 0* or *Sol 4*).

The results obtained demonstrated significant variations in the manner in which each system supported user adaptation, engagement, and trust. Systems such as `EZExFix` and `Harmony` demonstrated strong trajectories in experiential dimensions, including stimulation and trust. In contrast, the `Parrot` drone exhibited challenges in sustaining perceived usefulness and value over time. `Tap Strap 2` was distinguished by its versatility and psychological engagement, while `Romie` exhibited a discernible pattern of functional recovery following an initial adjustment period.

Interpretation of these results through the prism of established HCI frameworks, including Ability-Based Design, Reality-Based Interaction, and Sensorimotor Realities, enabled the contextualisation of user feedback within broader theoretical paradigms. Consequently, this facilitated the development of targeted design adaptation strategies, customised to the distinct constraints of extraterrestrial environments.

6.1 Future work

In this thesis, findings are presented that provide a foundation for ongoing research into the usability and adaptability of interactive systems in hostile environments. It is evident that several areas of potential research exist for future studies.

Firstly, future studies could benefit from broader and more detailed data collection, encompassing physiological metrics (e.g. stress levels, task load indices) alongside self-reported UX scores, which were only considered with EZExFix. This would facilitate a more comprehensive understanding of the manner in which cognitive, emotional, and environmental variables interact to influence system usability in space analogue missions.

Secondly, it is important to note that the present study concentrated on individual system evaluations in isolation. The exploration of inter-system interactions, such as the coordination between a wearable input device (Tap Strap 2) and a knowledge assistant (Harmony), would be a valuable extension of this research. The ability to comprehend system interoperability and ecosystem dynamics plays a pivotal role in the design of cohesive human-computer ecosystems that are suitable for long-duration missions.

Thirdly, while this study applied HCI frameworks retrospectively to analyse UX trends, future work could operationalise these frameworks proactively. One approach to achieving this would be to utilise Ability-Based Design or Sensorimotor Realities as a blueprint during system integration. This would facilitate the development of context-aware systems with the capacity for real-time personalisation in response to user constraints or mission demands.

Finally, as the scope of space travel and crew diversity expands, it must be noted that future UX evaluations should account for a broader spectrum of users, abilities, and cultural expectations. The integration of methods derived from user-centred design, particularly during the preliminary stages of mission preparation, has the potential to guarantee that the resulting systems are not only technically robust, but also inclusive and more psychologically supportive.

Bibliography

Allam, A. H., Hussin, A. R. C., & Dahlan, H. M. (n.d.). User Experience: Challenges and Opportunities.

Carroll, J. M. (2009). Human computer interaction (hci). *Encyclopedia of Human-Computer Interaction. Aarhus: The Interaction Design Foundation.*

Frost, J. (2022, July). Cronbach's Alpha: Definition, Calculations & Example. Retrieved May 24, 2025, from https://statisticsbyjim.com/basics/cronbachs-alpha/

Griffin, B. (2024, October). Calculating Cronbach's Alpha in Excel. Retrieved May 24, 2025, from https://psychologyofbusiness.beehiiv.com/p/calculating-cronbachs-alpha-excel

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, *25*(2), 91–97. https://doi.org/10.1080/01449290500330331

Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: A framework for post-WIMP interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 201–210. https://doi.org/10.1145/1357054.1357089

Jordan, P. W. (2000, April). *Designing pleasurable products.* CRC Press.

Karabacak, M., & Margetis, K. (2023). Embracing large language models for medical applications: Opportunities and challenges. *Cureus.* https://doi.org/10.7759/cureus.39305

Karapanos, E., Hassenzahl, M., & Martens, J.-B. (2008). User Experience Over Time. *Works In Progress.*

Law, L.-C., Roto, V., Hassenzahl, M., Vermeeren, A., & Kort, J. (2009). Understanding, scoping and defining user experience: A survey approach. *Proc. CHI '09*, 719–728. https://doi.org/10.1145/1518701.1518813

M.A.R.S. UCLouvain - Martian simulation. (n.d.). Retrieved May 21, 2025, from https://www.marsuclouvain.be/

Machno, A., Jannin, P., Dameron, O., Korb, W., Scheuermann, G., & Meixensberger, J. (2015). Ontology for assessment studies of human–computer-interaction in surgery. *Artificial Intelligence in Medicine*, *63*, 73–84. https://doi.org/10.1016/j.artmed.2014.12.011

Manon, J., Englebert, A., & Cornu, O. (2023). Le traitement des fractures en environnement hostile: De l'espace à la médecine de guerre... tout le monde peut-il le faire? *Ortho-Rhumato*, *21*, 6.

Mrazek, K., Holton, B., Klein, T., Khan, I., Ayele, T., & Khan Mohd, T. (2021). The tap strap 2: Evaluating performance of one-handed wearable keyboard and mouse. *HCI International 2021-Late Breaking Papers: Multimodality, eXtended Reality, and Artificial Intelligence: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings 23*, 82–95.

Pataranutaporn, P., Sumini, V., Ekblaw, A., Yashar, M., Häuplik-Meusburger, S., Testa, S., Obrist, M., Donoviel, D., Paradiso, J., & Maes, P. (2021). Spacechi: Designing human-computer interaction systems for space exploration. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411763.3441358

Pataranutaporn, P., Sumini, V., Yashar, M., Testa, S., Obrist, M., Davidoff, S., Paul, A. M., Donoviel, D., Wu, J., Fish, S. A., Ekblaw, A., Schmidt, A., Paradiso, J., & Maes, P. (2022). SpaceCHI 2.0: Advancing Human-Computer Interaction Systems for Space Exploration. *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. https://doi.org/10.1145/3491101.3503708

Review: Hands On With the Parrot ANAFI - DRONELIFE. (n.d.). Retrieved May 27, 2025, from https://dronelife.com/2018/07/24/review-drone-parrot-anafi/

Saint-Guillain, M., Vanderdonckt, J., Burny, N., Pletser, V., Vaquero, T., Chien, S., Karl, A., Marquez, J., Wain, C., Comein, A., Casla, I. S., Jacobs, J., Meert, J., Chamart, C., Drouet, S., & Manon, J. (2023). Enabling astronaut self-scheduling using a robust advanced modelling and scheduling system: An assessment during a Mars analogue mission. *Advances in Space Research*, *72*(4), 1378–1398. https://doi.org/https://doi.org/10.1016/j.asr.2023.03.045

Sarumi, O. A., & Heider, D. (2024). Large language models and their applications in bioinformatics. *Computational and Structural Biotechnology Journal*. https://doi.org/10.1016/j.csbj.2024.09.031

Schrepp, M., & Thomaschewski, J. (n.d.). Design and Validation of a Framework for the Creation of User Experience Questionnaires. *International Journal of Interactive Multimedia and Artificial Intelligence*, *5*.

Tribolet, A., & Officer, C. G. (n.d.). Exploring Drone Utility for Radiation Protection on Martian Ground and Underground.

Vanderdonckt, J., Vatavu, R.-D., Manon, J., Saint-Guillain, M., Lefevre, P., & Marquez, J. J. (2024). Might as Well Be on Mars: Insights on the Extraterrestrial Applicability of Interaction Design Frameworks from Earth. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3613905.3650807

Vatavu, R.-D. (2022). Sensorimotor Realities: Formalizing Ability-Mediating Design for Computer-Mediated Reality Environments. *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 685–694. https://doi.org/10.1109/ISMAR55827.2022.00086

Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-Based Design: Concept, Principles and Examples. *ACM Transactions on Accessible Computing*, *3*(3), 1–27. https://doi.org/10.1145/1952383.1952384

Analysis of User Experience of Interactive Applications by Analog Astronauts

This thesis is part of a wider study of Human-Computer Interactions (HCIs) (HCIs) in the context of the conquest of space. The aim of this study was to analyse the results of user experiments involving five different devices utilised as part of the M.A.R.S. UCLouvain mission. The devices included Romie (an advanced modelling and scheduling management application), EZExfix (an external fixator for bone fractures), Tap Strap 2 (a single-handed, all-in-one wearable gesture controller), Harmony (an LLM designed to assist astronauts), and a Parrot drone. This thesis establishes the foundations for improving the adaptability and usability of interactive systems in hostile environments. To achieve this, data collection must be broadened to encompass both physiological metrics and user experience (UX), with the aim of investigating interactions between systems. Furthermore, HCI frameworks should be integrated into the product design process. Attention must also be paid to user diversity in order to create inclusive systems suitable for future space missions.

Ce travail s'inscrit dans le cadre d'une étude plus large sur les interactions homme-machine (ICH) dans le contexte de la conquête spatiale. L'objectif de cette étude était d'analyser les résultats d'expériences d'utilisateurs impliquant cinq dispositifs différents utilisés dans le cadre de la mission M.A.R.S. UCLouvain. Les dispositifs comprenaient Romie (une application avancée de modélisation et de gestion de planning), EZExfix (un fixateur externe pour les fractures osseuses), Tap Strap 2 (un contrôleur gestuel portable tout-en-un pour une seule main), Harmony (un LLM conçu pour assister les astronautes), et un drone Parrot. Ce travail établit les bases pour améliorer l'adaptabilité et l'utilisabilité des systèmes interactifs dans des environnements hostiles. Pour ce faire, la collecte de données doit être élargie pour englober à la fois les mesures physiologiques et l'expérience utilisateur (UX), dans le but d'étudier les interactions entre les systèmes. En outre, les cadres de l'IHM devraient être intégrés dans le processus de conception des produits. Il convient également de prêter attention à la diversité des utilisateurs afin de créer des systèmes inclusifs adaptés aux futures missions spatiales.

Dylan Goffinet, Pierre Huljev

Université catholique de Louvain
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique
www.uclouvain.be/lsm