Decoding Auditory Feedback: Enhancing Usability of RFID and QR code scanning methods with Sound Insights

Lode Dams
Rentman
Utrecht, The Netherlands
lode@rentman.nl

Christof van Nimwegen
Utrecht University
Department of Information and Computing Sciences
Utrecht, The Netherlands
c.vannimwegen@uu.nl

Frans Wiering
Utrecht University
Department of Information and Computing Sciences
Utrecht, The Netherlands
f.wiering@uu.nl

ABSTRACT

This study compares the impact of *conventional auditory earcons* and *enhanced auditory earcons* on the usability of handheld devices for different scanning methods in Warehouse Management Systems (WMSs). Warehouses deal with labor shortages and optimization challenges. There is a growing recognition of the significance of incorporating multimodality to craft immersive user experiences (UX) and to provide natural and robust interaction. While visual feedback is commonly used, audio feedback during QR code and RFID scanning is often limited to auditory earcons consisting of beeps. Our experimental research includes qualitative and quantitative measures exploring, verifying, and validating *enhanced auditory earcons* for WMSs. During the evaluation phase, four conditions are compared on task completion time, number of errors, perceived workload, annoyance, and perceived usability. *Enhanced auditory feedback* yielded lower perceived workload, was less frustrating and less annoying compared to *conventional earcons*. RFID scanning proved to be more efficient and effective, while QR code scanning was less mentally demanding. These results contribute to a better understanding of how usability in warehouse ict-artefacts could be improved, and could be extended to other domains such as retail or transport.

CCS CONCEPTS

• **Human-centered computing → User studies**.

KEYWORDS

Auditory icons, earcons, auditory feedback, warehouse management system, RFID, QR, usability

ACM Reference Format:

Lode Dams, Christof van Nimwegen, and Frans Wiering. 2024. Decoding Auditory Feedback: Enhancing Usability of RFID and QR code scanning methods with Sound Insights. In *European Conference on Cognitive Ergonomics (ECCE 2024), October 08–11, 2024, Paris, France.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3673805.3673822

ECCE 2024, October 08–11, 2024, Paris, France

ACM ISBN 979-8-4007-1824-3/24/10
https://doi.org/10.1145/3673805.3673822

1 INTRODUCTION

User interfaces (UIs) traditionally are designed mainly for the human visual system. However, the importance of multimodality in creating immersive and engaging user experiences (UX) is nowadays recognized as we are moving towards multimodal user interfaces [6, 20]. Overly dense visual displays can lead to cognitive overload, negatively impacting the user's performance. Past studies have shown that information spread over more than one modality helps to reduce the user's cognitive load, is perceived as more usable and enjoyable [1]. More specifically, audio has proven to give valuable feedback to users' actions, carry information, provide information beyond the field of vision, enhances visual representation, and can strengthen the emotion and immersion that a UI creates [13]. Different types of auditory feedback can be distinguished, see Figure 1. Firstly, *auditory icons* are familiar sounds based on experiences in the real world, e.g. the sound of paper crumbled up when an item is moved to the trash can [11]. Secondly, *earcons* are abstract, synthetic, and musical tones or sound patterns which have a metaphorical relation with the object they represent. Blattner et al. [5] define earcons as "audio messages that are used in the computer/user interface to provide information about some computer object, operation or action". An example of an earcon is the initiation sound of Apple's Siri.

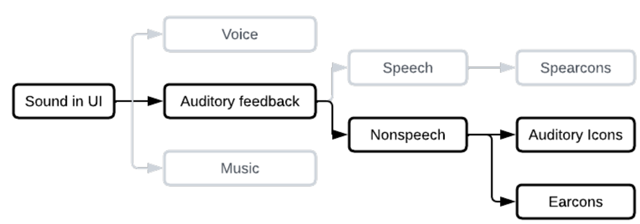

Figure 1: Scope of this study.

Having been neglected for long, the use of audio in UIs is now receiving more attention in e.g. the video game industry, retail, and automotive domain [15]. The intricate relationship between sounds and emotions forms a cornerstone of human experience. Our auditory sense plays a vital role in our cognitive and emotional processes, serving multiple functions. It acts as a vigilant guardian, constantly gathering and processing acoustic information to ensure our safety and survival. Simultaneously, our ability to hear enriches our emotional landscape, adding depth and nuance to our daily interactions with the world around us. The connection between auditory input and emotional output underscores the significance of sound in the human experience, highlighting its power to influence our moods, memories, and overall well-being. However, despite advancements in audio technology and lots of research, many UIs still fail to fully leverage the potential of sound, leading to UX related emotions such as frustration, incomprehension, and annoyance [9]. There is no clear heuristic guidance for sound when designing interfaces [20]. One domain needing additional research to fully leverage the potential of auditory feedback is the Warehouse Management Systems (WMSs) domain. While visual feedback during scan processes in mobile WMS applications is typically provided through a visual display, audio feedback is often overlooked or limited to a beep by the scanner. Enhanced auditory feedback (e.g. auditory icons and earcons) during scan processes should be considered in these WMSs [13], which effectively changed how many employees do their jobs. One of the changes by technological advancements is RFID (Radio Frequency Identification) and QR code scanning [22], where this study focuses on. The following research question will be answered:

How do conventional auditory earcons and enhanced auditory earcons impact the usability of different scanning methods in Warehouse Management Systems?

In this study an experiment is conducted where participants test both conventional auditory earcons and enhanced auditory earcons for both QR code and RFID scanning. Results of this will contribute to knowledge about the usability of auditory feedback with the relatively new RFID technology and already existing QR code technology in warehouses.

2 RELATED WORK

2.1 The role of auditory feedback in UI

There have been strides in the domain of auditory feedback in wearables and hand-held devices such as smartwatches, tablets and mobile phones [1, 20]. Non-speech sounds are used in audio-tactile feedback on touch interfaces, swiping a page in an e-reader, exceeding the speed limit detected by a mobile GPS, etc. The role of auditory feedback is becoming increasingly important for conveying information, enhancing UX, and improving accessibility.

2.1.1 Conveying information. Vision is often considered the most valued sense while hearing is ranked second [5], but due to limited screen space there is the risk of overly dense displays leading to cognitive overload, negatively impacting performance [5, 20]. It showed that information spread over multiple modalities helps to minimize users' cognitive load and increases task performance [1]. Displaying information on the auditory channel is useful because it enables receiving immediate and informative feedback, without solely relying on visual elements. Aside from relieving the visual channel, audio can be useful when the user cannot view a display (e.g. while driving), and it be heard from 360 degrees and over distance[7, 13]. However, it cannot always be replayed, unlike an icon it is temporary, potentially resulting in information loss. Compared to visual stimuli auditory stimuli evoke a faster reaction time, but perceptual judgments by the eyes are usually more precise than those made with the ears. Still the under-utilized auditory channel can play a significant role and offer advantages in situations where visual attention is limited. The auditory modality offers valuable opportunities to improve user interfaces and optimize information transfer. By using the strengths of both visual and auditory channels, designers can create interfaces with increased usability.

2.1.2 User Experience. Multimodality is assumed to provide a more natural and robust interaction than unimodal systems and consequently enhance usability and UX. Besides positive effects of multimodality on cognitive load, multimodal interfaces are perceived to be more usable and enjoyable, where audio is believed to increase stimulation and pleasurable experiences [1]. It was also found that melody-based feedback rather than beeps, improves effectiveness, efficiency, and user satisfaction [17]. It showed that combining audio with graphics significantly enhances usability by leveraging our ability to process information across senses. [6]. Although guidelines exist, [9] claim that designing auditory feedback in UIs is still poorly understood. A major UX concern is users quickly linking auditory feedback to annoyance[6]. Future research should focus on mitigating perceived annoyance and improving user satisfaction.

2.2 Auditory icons

Auditory icons are everyday sounds meant to convey information about events by analogy with everyday events [10] based on experiences in the real world. Gaver [10] claims that auditory icons are beneficiary as they is based on the way people normally listen. Auditory icons represent information in an intuitive way, providing information that visual displays do not, and auditory icons and visual icons together create a more encompassing world for the user. There are factors that affect the usability of both auditory icons and earcons: 1) *Meaningfulness*: the relation between the auditory feedback and referent, 2) *Learnability*: how easily auditory feedback can be learned, 3) *Identification*: the extent to which auditory feedback can be easily perceived and separated from other cues, 4) *Musical characteristics*: the characteristics belonging to well-designed auditory feedback, and 5) *User preferences*: opinions of users on the auditory feedback.

Although not widely employed yet, auditory displays equipped with both auditory icons and earcons are becoming increasingly prominent in home appliances, computers, smartphones, automotive, aviation, medical, financial, and military applications [20]. In the automotive domain, auditory icons inform drivers about the condition of the vehicle, prevent misbehavior such as falling asleep behind the wheel or forgetting to lock the doors, and improve situational awareness by reducing brake response time [3]. Although the importance of sound is well-known in video games, there is a growing interest in using auditory icons as means of providing the player with *additional* information [16]. Auditory icons providing multidimensional information are promising as they convey many different attributes of their source [16]. Airplane pilots have to deal with a great amount of information, including a visually demanding interface. A study by Perry et al. [18] investigated the efficacy of auditory icons as warning signals in an aviation context. Significantly fewer training trials were required to learn auditory icon warnings compared to conventional warnings, and accuracy in the test phase was higher for auditory icons. Conventional warnings elicited slow reaction times and poor accuracy. These findings can help pilots to learn the interface quicker and make fewer mistakes. Studies on hospitals and medical equipment have reported similar results regarding learnability too . Lastly, auditory icons also have proven to positively impact notifications on mobile devices. Learnability, memorability, and intuitiveness are better when auditory icons are used for notifications.

2.3 Earcons

Auditory earcons [21], unlike auditory icons with direct analogies are abstract, e.g. a sound when pressing a key on a touchscreen. They can range from from basic beeps signaling errors to elaborate auditory themes (e.g. in contemporary OSs, denoting actions like startup and various other events). Blattner et al. [5] discuss several types of earcons:

(1) *One-element earcons*: Single-pitch earcons are used to transmit a single bit of information (e.g. saving, clicking), they cannot be decomposed further [5]. Single-motive earcons on the other hand have a brief succession of pitches to produce a rhythmic and tonal pattern sufficiently distinct to allow it to function as an individual, recognizable entity. The attributes included are rhythm, pitch, timbre, register, and dynamics. They represent common computer entities such as error messages, system information, windows, and files [5].
(2) *Compound earcons* are formed by placing two or more one-element earcons in succession. If a single-pitch earcon is created for the icon "file" and a single-pitch earcon is created for the action "open", then a compound earcon is "open file" by placing two single-pitch earcons after each other (figure 2).
(3) *Hierarchical earcons* are constructed around "grammar", where each earcon is a branch of a tree and each branch receives all the properties of the branches above it in the tree. Hierarchical earcons are useful for systems with a large number of earcons present as the message can become very sophisticated [5], see Figure 3.

Figure 2: Compound earcons by Roginska [20].

2.4 Applications

Certain domains, such as the automotive domain benefit from earcons in many ways. A study by Monsaingeon et al. [15] showed that earcons were efficiently perceived and provoked a small decrement to a visual task, meaning the driver no longer has to look at his display as often. A study by Reynolds-McIlnay and Morrin

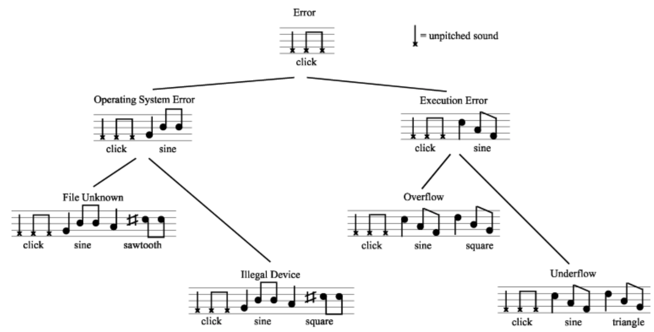

Figure 3: Hierarchical earcons by Blattner et al. [5].

[19] discussed the importance of Retail Transaction Auditory Confirmation (RTAC) providing earcons during purchase transactions. Purchase transactions happen in a visual and auditory complex environment, having a distracting nature as one could potentially interact with other individuals. RTAC enhances trust by utilizing earcons associated with purchase transactions, such as the beep heard when scanning an item at the checkout. These earcons offer confirmation that the technology has successfully registered their actions, minimizing uncertainty during the transaction process. Like auditory icons earcons are used in games too. Earcons have the advantage of being context-free and thus can represent any event or interaction in the interface. They also tend to be more precise than auditory icons. However, the disadvantage is that earcons have no intuitive knowledge to draw on when interpreting them; they have to be learned [16].

2.5 Comparison of auditory icons and earcons

Several studies suggest that auditory icons can be easier to learn and that reaction times to auditory icons can be shorter than to earcons, while other studies have demonstrated that earcons can be more pleasant in certain cases [2]. It might also be that users' retention of sounds depends heavily on the individual sound (not the sound type) as well as the learning method used. It is certain that auditory icons and earcons evoke different kinds of cognitive capabilities. For auditory icons, this means they are easier associated with iconic entities, while earcons are used where no reference to a physical entity is available [8]. Several researchers have argued that this strict discrimination between auditory icons and earcons may be limiting for real-life applications, as no clear qualitative dominant superiority has been found. They suggest that both types should be used together, while others also highlight that auditory icons and earcons are theoretical extremes along a continuum of semi-abstract non-speech sounds [8].

2.6 Audio for Mobile Warehouse Management Systems

A WMS is an important aspect of a supply chain network as it aims to control the movement and storage of materials within a warehouse and process the associated transactions, including shipping, receiving, put-away, and picking. It is the interface used to manage processes, people, and equipment on the operational level. Both efficiency and effectiveness are crucial factors for WMSs, as companies always strive for minimizing warehousing costs and increasing throughput rates. WMSs often use Auto ID Data Capture (AIDC) technology such as as QR code or RFID scanners. Within a warehouse, pickers pack projects with materials equipped with QR codes or RFID tags by using a scanning device, see Figure 4. Auditory feedback in the form of beeps is used to confirm a scan is performed. As soon as a QR code or RFID tag is read, the material is packed and processed.

Figure 4: The Zebra RFD40 which can scan both RFID tags and QR codes, and examples of an RFID-tag and QR code.

There are active, passive, and semi-passive RFID tags. Passive tags play a vital role in WMSs because of their small size, low power consumption, low cost, robustness, and little interference [22]. Passive tags do not require a power source, as the energy is transferred from the reader to the tag. Semi-passive and active tags require built-in batteries, are larger in size, and are more difficult to handle. This study assumes passive tags are used as we focus on the warehousing domain. See figure 4 for a RFID tag. RFID could be seen as a replacement for QR code technology. Although QR code technology is low-priced, compact, and has low power consumption, it still needs a direct line of sight and is susceptible to light sources. RFID tags have more data capacity storage and are not dependent on undamaged labels. On top of that, multiple RFID tags can be scanned at once, while QR codes can only be scanned consecutively [22]. Although AIDC in the form of RQ code and RFID scanning is widely accepted and used, the latter is rarely covered in warehouse-related research publications. Additionally, auditory feedback for both QR code and RFID scanning is still relatively untouched. In current practice users heavily rely on the familiar beep sound as confirmation of a scan. While research has shown that feedback such as auditory icons or earcons can enhance performance and reduce errors, only a few studies have explored auditory feedback during scan processes [4]. The importance of improved auditory feedback is evident in other domains as well, e.g. in the healthcare sector, bedside QR code scanning systems rely on auditory beeps for medication verification. However, the use of identical beeps for correct and incorrect scans can lead to confusion among nurses, who may mistakenly assume that the correct medication has been scanned. The retail domain showed that auditory confirmation with a beep enhanced trust, reduced cognitive load, and positively impacted customer satisfaction [19]. Drawing from these findings, the application of more comprehensive auditory feedback, beyond a simple beep, during the pick-by-scan process in the warehouse domain could potentially improve usability.

3 METHOD

This study assessed the usability of handheld devices for both QR code and RFID scanning by analyzing the impact of using enhanced auditory earcons instead of conventional earcons (simple beeps), with the following research question:

How do conventional auditory earcons and enhanced auditory earcons impact the usability of different scanning methods in Warehouse Management Systems?

The research method combined qualitative data about opinions and perceptions, and quantitative data about performance to provide a comprehensive understanding of the impact different types of auditory feedback on usability in WMSs. In an experiment the usability of auditory feedback was assessed by researching the effectiveness, efficiency, perceived workload, annoyance, and perceived usability in different conditions. Perceived workload was be established by using the NASA-TLX [12], perceived usability was be measured with the SUS [14]. The following sub-questions were formulated:

SQ1: *How do conventional and enhanced auditory earcons affect efficiency during the scan process?*
SQ2: *How do conventional and enhanced auditory earcons feedback affect effectiveness during the scan process ?*
SQ3: *How do conventional and enhanced auditory earcons affect cognitive workload during the scan process?*
SQ4: *How do conventional and enhanced auditory earcons affect annoyance during the scan process?*
SQ5: *How do conventional and enhanced auditory earcons affect the perceived usability of the scan process?*

3.1 Experimental design

3.1.1 Independent variables. The independent variables were **Scan type** (QR code vs. RFID) and **Sound type** (conventional auditory earcons vs. enhanced auditory earcons). Combined they led to a 2x2 factorial design with 4 conditions:

(1) **QC**: QR code scanning with Conventional auditory earcons
(2) **QE**: QR code scanning with Enhanced auditory earcons
(3) **RC**: RFID scanning with Conventional auditory earcons
(4) **RE**: RFID scanning with Enhanced auditory earcons

3.1.2 Dependent variables. The five dependent variables were:

(1) Efficiency (task completion time in seconds)
(2) Effectiveness (number of errors made)
(3) Perceived workload (NASA-TLX, questions scored on a 20-item likert-scale, multiplied by 5 to obtain 0-100 score)
(4) Perceived annoyance (question scored on a 20-item likert-scale, multiplied by 5 to obtain 0-100 score)
(5) Perceived usability (SUS, statements rated on a 5-item likert-scale which after a formula results in a 0-100 score)

3.1.3 Participants. Participants for the experiment were selected using a convenience sampling procedure. 41 participants were recruited from a company involved (36) and among university students (5). All were adults, of which 27 were male (66%) and 14 were female (35%), with a mean age of 29.7 years (*SD* = 6.7).

3.1.4 Materials. In the experiment 30 items with QR codes *or* RFID tags had to be scanned. The scanner used was the Zebra RFD40 for both QR codes and RFID tags. The sounds that were used:

- The audio in the *conventional auditory earcons* conditions were one-element earcons (see section 2.3), they were default sounds and consisted solely of beeps
- The audio used in the *enhanced auditory earcons* conditions were new. These enhanced auditory earcons have been carefully designed in a pre-study in an iterative manner, with in-between-evaluations. The enhanced auditory icons were a combination of Compound earcons and Hierarchical earcons (see section 2.3)

3.1.5 Setup and procedure. The experiment was conducted in a quiet environment where one could walk around to simulate the pick-by-scan process. Participants were acquainted with the WMS interface and briefly practiced scanning products, after which they performed the real scanning tasks in the four different conditions. Participants underwent the scanning tasks in four different conditions, which were counterbalanced to minimize fatigue and learning effects. They had to scan 30 items and move them from location A to B. Error messages were included to improve realism and test auditory feedback for them. After this participants answered the Annoyance question and filled in the NASA-TLX and the SUS questionnaires.

4 RESULTS

We here elaborate on the results, which are also listed in Table 1 at the end of this section.

4.1 Efficiency

The measure for efficiency was *time (s)*. Note that the focus of this study is primarily on sound, not on the difference between scan methods in itself. RFID tags are read by radio signal, whereas QR codes are optical constructs, and therefore have to be read by an optical device. RFID tag reading was expected to be faster because they have a significantly longer range. This also showed here. Assumptions for normality and homogeneity were met. ANOVA showed a main effect for *scan type*, $F(1, 37) = 15.78$, $p < .001$, RFID scanning ($M = 173.0$, $SD = 35.4$) was significantly faster than QR scanning ($M = 186.8$, $SD = 33.4$). On average, RFID scanning was 7.3% faster compared to QR scanning. There was no effect (nor interaction effects) of *sound type*; enhanced auditory earcons did not impact the work pace.

4.2 Effectiveness

The measure for effectiveness was *errors*. Assumptions for normality and homogeneity were *not* met, so a non-parametric two-sided Wilcoxon signed rank was used. Again a main effect for *scan type* was found, RFID scanning ($M = 0.52$, $SD = 0.73$) caused less errors than QR scanning ($M = 1.02$, $SD = 1.18$), $T = 472.0$, $Z = -5.68$, $p = .004$. RFID scanning was 51.0% more accurate than QR scanning. There was again no effect of *sound type* (nor interaction effects); enhanced auditory earcons did not impact the effectiveness of the employee.

4.3 Annoyance

The annoyance measure was the score on one question: "*How annoyed were you?*" scored on a 20-item likert-scale, multiplied by 5 to obtain a 0-100 score. Assumptions for normality and homogeneity were *not* met, and a non-parametric two-sided Wilcoxon signed rank was used. For *scan type*, no effects were found. However, *sound type* yielded a significant difference, $T = 801.50$, $Z = -4.16$, $p = .023$. Enhanced auditory earcons (M = 28.8, SD = 24.6) was perceived as less annoying than conventional earcons (M = 33.6, SD = 24.4).

4.4 Perceived usability

The measure for Perceived usability was the SUS score. Assumptions for normality and homogeneity were met, therefore a two-way repeated measures ANOVA was performed. There was a significant effect of *scan type*, $F(1, 37) = 10.92$, $p = .002$. QR scanning ($M = 84.1$, $SD = 12.9$) scored higher on the SUS than RFID scanning ($M = 77.1$, $SD = 17.5$), see Figure 5. *Sound design* had no effects on the SUS-scores.

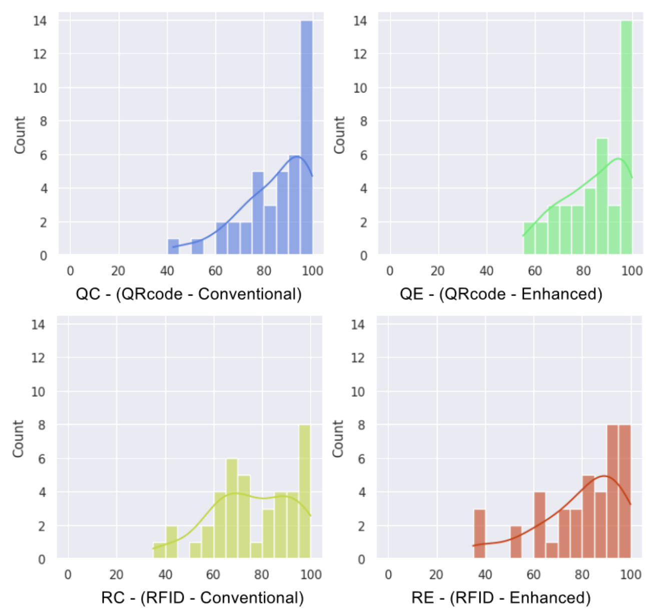

Figure 5: Distribution of the SUS-scores by condition

4.5 Perceived workload

The measure for perceived workload efficiency was the TLX score, which was calculated by taking the average of the six NASA-TLX dimensions. Since assumptions for normality and homogeneity were met, a two-way repeated measures ANOVA was used. This time *Sound type* yielded a main effect, $F(1, 37) = 4.37$, $p = .043$, indicating a significant difference between conventional earcons (M = 30.0, SD = 16.5) and enhanced auditory earcons (M = 27.6, SD = 14.1). The perceived workload of the enhanced auditory feedback was 2.4 points lower compared to the conventional earcons, meaning it is less demanding. There were no effect (nor interaction effects) of *scan type*.

Besides the calculated TLX average, tests were performed on individual NASA-TLX factors for more insight. Two statistically significant effects were found: there was a main effect of the *scan type* on mental demand, $F(1, 37) = 7.73$, $p = .008$. QR code scanning (M = 24.3, SD = 20.4) was significantly less mentally demanding than RFID tag scanning (M = 31.0, SD = 23.4). Besides, there was a main effect of *sound type* on frustration, $F(1, 37) = 4.24$, $p = .046$. Enhanced auditory earcons (M = 24.8, SD = 20.6) was less frustrating than conventional sounds (M = 29.5, SD = 22.4).

Table 1: Resume of significant results of experiment

Dep.	Indep.	Result	Comparison
Efficiency	Scan	RFID scanning is faster than QR code scanning	173s vs. 187s (7.3%)
Effectiveness	Scan	RFID scanning has fewer errors compared to QR code scanning	0.52 vs. 1.02 (51.0%)
Workload	Sound	Enhanced auditory feedback has lower perceived workload than conventional sounds	27.6 vs. 30.0 (8.0%)
Workload	Scan	QR code scanning requires less mental demand than RFID scanning	24.3 vs. 31.0 (21.6%)
Workload	Sound	Enhanced auditory feedback is less frustrating than conventional sounds	24.8 vs. 29.5 (15.9%)
Annoyance	Sound	Enhanced auditory feedback is less annoying than conventional sounds	28.8 vs. 33.6 (14.3%)
SUS	Scan	QR code scanning has higher usability score than RFID scanning	84.1 vs. 77.1 (9.1%)

5 DISCUSSION

5.1 Influence of scan type on usability

RFID scanning proved to be faster than QR scanning. This might somehow be expected, since RFID does not need a direct line of sight and multiple RFID tags can be scanned at once. RFID scanning was 7.3% faster and therefore more efficient. Also regarding effectiveness (expressed by errors made) RFID performed better as less errors were made here. This could be seen as surprising, as one would think that scanning items one by one with QR code scanning gives a user more control about what has been scanned and what has not. However, many participants double scanned QR codes during the experiment, something which does not happen as frequently with RFID due to the nature of RFID scanning.

Two significant results were in favor of QR scanning: QR code scanning was perceived as less mentally demanding than RFID scanning. One of the reasons might be the fact that people are more used to QR code scanning in their daily life. Lastly, QR scanning yielded a higher SUS score than RFID scanning. Both scan types score above 68, which is above average. While RFID falls in the B category (68-80.3), QR scanning is in A (>80.3), meaning the latter is in the highest usability segment. Resuming, while RFID scanning was shown to be more effective and efficient, in terms of mental demand and usability QR code scanning is preferred which is probably because of its direct and simplistic nature.

5.2 Influence of sound type on usability

In terms of perceived workload, it showed that *enhanced auditory earcons* outperform *conventional auditory earcons* (just beeps), for both QR code and RFID scanning. A lower perceived workload could potentially be beneficial for increased pleasure at work and less stress. *Enhanced auditory earcons* primarily focused on improving the three roles of auditory feedback in UI (conveying information, enhancing UX, and improving accessibility). It seems safe to say that it is useful to implement *enhanced auditory earcons* for users to have a more enjoyable experience and experience less workload. *Enhanced auditory earcons* also proved to be less frustrating and annoying than *conventional auditory earcons*. For systems where sounds are heard over a thousand times a day this could be important. Having significantly less frustration and being significantly less annoyed increases UX, it keeps employees more in the flow, and can be potentially beneficial for long-term efficiency or even employee happiness.

5.3 Limitations

There could be limitations impacting the interpretation and generalization of the findings. The participant pool primarily consisted of employees of the collaborating company, who do not represent *a wide range* of potential end users. Performing the experiment at different end user warehouses would introduce confounding factors such as background noises and setup variations, compromising the reliability and validity of the results. Also, there was a learning effect during the experiment; on average, participants became faster at their tasks as they learned where the QR code sticker was, became better at RFID scanning, and improved their efficiency by picking up items and moving them from A to B.

6 CONCLUSION AND FUTURE WORK

Several directions for future research can be identified. Our study focused solely on the influence of enhanced auditory earcons on top of visual feedback. Future research could explore the integration of haptic / tactile feedback to assess how a holistic multimodal approach enhances usability. As virtual and augmented reality technologies become increasingly prevalent, enhanced auditory earcons could play a pivotal role in spatial awareness and immersion. Regarding ecological validity, certainly conducting studies in real-world warehouse environments where background noise is present should take place, as well longitudinal investigation into how users' experiences evolve over an extended period of system usage are useful (since our study captured only a short timeframe). Comparative analyses with other industries that employ similar scanning and feedback systems could be performed. This might provide valuable benchmarks to help identify best practices, lessons learned, and potential transferable insights for optimizing usability in not only the warehousing domain, but also domains such as retail, transport, and healthcare. Lastly, as technology continues to advance, one can perhaps anticipate the integration of AI-driven algorithms for context-aware audio cues, fostering a more intuitive and immersive interaction between users and digital environments.

REFERENCES

[1] Rafa Absar and Catherine Guastavino. 2015. The design and formative evaluation of nonspeech auditory feedback for an information system. *Journal of the Association for Information Science and Technology* 66, 8 (2015), 1696–1708.

[2] T S Amer and Todd L. Johnson. 2018. Earcons versus auditory icons in communicating computing events: learning and user preference. *International Journal of Technology and Human Interaction (IJTHI)* 14, 4 (2018), 95–109.

[3] Steven M. Belz, Gary S. Robinson, and John G. Casali. 1999. A new class of auditory warning signals for complex systems: Auditory icons. *Human factors* 41, 4 (1999), 608–618.

[4] Samuel M. Berger and Timothy D. Ludwig. 2007. Reducing warehouse employee errors using voice-assisted technology that provided immediate feedback. *Journal of Organizational Behavior Management* 27, 1 (2007), 1–31.

[5] Meera M. Blattner, Denise A. Sumikawa, and Robert M. Greenberg. 1989. Earcons and icons: Their structure and common design principles. *Human–Computer Interaction* 4, 1 (1989), 11–44.

[6] Stephen A. Brewster and Murray G. Crease. 1999. Correcting menu usability problems with sound. *Behaviour & Information Technology* 18, 3 (1999), 165–177.

[7] Stephen A. Brewster, Peter C. Wright, and Alistair D.N. Edwards. 1993. An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. 222–227.

[8] Ádám Csapó and György Wersényi. 2013. Overview of auditory representations in human-machine interfaces. *ACM Computing Surveys (CSUR)* 46, 2 (2013), 1–23.

[9] Christopher Frauenberger, Tony Stockman, and Marie-Luce Bourguet. 2007. A survey on common practice in designing audio in the user interface. In *Proceedings of HCI 2007 The 21st British HCI Group Annual Conference University of Lancaster, UK 21*. 1–9.

[10] William W. Gaver. 1987. Auditory icons: Using sound in computer interfaces. *ACM SIGCHI Bulletin* 19, 1 (1987), 74.

[11] Richard Gould. 2016. Auditory Icons. https://designingsound.org/2016/11/24/auditory-icons/

[12] Sandra G. Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[13] Hannu Korhonen, Jukka Holm, and Mikko Heikkinen. 2007. Utilizing sound effects in mobile user interface design. In *Human-Computer Interaction–INTERACT 2007: 11th IFIP TC 13 International Conference, Rio de Janeiro, Brazil, September 10-14, 2007, Proceedings, Part I 11*. Springer, 283–296.

[14] Page Laubheimer. 2018. Beyond the NPS: Measuring Perceived Usability with the SUS, NASA-TLX, and the Single Ease Question After Tasks and Usability Tests. https://www.nngroup.com/articles/measuring-perceived-usability/

[15] Noé Monsaingeon, Loïc Caroux, Sabine Langlois, and Céline Lemercier. 2023. Earcons to reduce mode confusions in partially automated vehicles: Development and application of an evaluation method. *International Journal of Human-Computer Studies* 176 (2023), 103044.

[16] Patrick Ng and Keith Nesbitt. 2013. Informative sound design in video games. In *Proceedings of the 9th Australasian conference on interactive entertainment: matters of life and death*. 1–9.

[17] Minyong Park. 2011. Evaluation of earcons for portable digital electronic products considering perceptual user characteristics. *Human Factors and Ergonomics in Manufacturing & Service Industries* 21, 5 (2011), 464–473.

[18] Nathan C. Perry, Catherine J. Stevens, Mark W. Wiggins, and Clare E. Howell. 2007. Cough once for danger: Icons versus abstract warnings as informative alerts in civil aviation. *Human factors* 49, 6 (2007), 1061–1071.

[19] Ryann Reynolds-McIlnay and Maureen Morrin. 2019. Increasing shopper trust in retailer technological interfaces via auditory confirmation. *Journal of Retailing* 95, 4 (2019), 128–142.

[20] Agnieszka Roginska. 2013. Auditory icons, earcons, and displays: Information and expression through sound. *The psychology of music in multimedia* 339 (2013).

[21] Denise A. Sumikawa. 1985. *Guidelines for the integration of audio cues into computer user interfaces.* Technical Report. Lawrence Livermore National Lab., CA (USA).

[22] B.S.S. Tejesh and S. Neeraja. 2018. Warehouse inventory management system using IoT and open source framework. *Alexandria engineering journal* 57, 4 (2018), 3817–3823.