# Holistic Multimodal Interaction and Design

Eric Chan, Gerry Chan, Assem Kroma, and Ali Arya(✉)

School of Information Technology, Carleton University, Ottawa, Canada
{erics.chan,gerry.chan,assem.kroma,ali.arya}@carleton.ca

**Abstract.** This paper addresses the lack of an HCI framework that examines multimodal experiences in a holistic way to help design interactions where modalities are properly integrated. Such a holistic approach is important as new emerging technologies and experiences are being introduced and used frequently. The resulting multimodal experiences present new requirements and challenges which are not managed by existing HCI frameworks. Through a literature review, we propose a set of principles that define the Holistic Multimodal Interaction and Design (HMID) framework. We perform a user study as the initial evaluation of the effectiveness of this framework. The study findings showed the potential value of HMID and suggested improvements to its guiding principles.

**Keywords:** Multimodal systems · Human-computer interaction · Holism · Retail

## 1 Introduction

Human-Computer Interaction (HCI) has witnessed a trend from systems based on a single modality of interaction (i.e., method, sense, or mode of operation used to do or experience something) to multimodal ones where the user can choose which modality suits their tasks better. Emerging technologies such as the Internet of Things (IoT), wearables, and Augmented/Virtual Reality (AR/VR) have accelerated this multimodal interaction trend. Multimodal interfaces enable the user to employ different modalities such as voice, gesture, and typing for communicating with a computer [1, 2]. Various HCI frameworks such as WIMP [3], Tangible [4], Ubiquitous [5], Multimodal [1, 2], and Natural [6] have offered increasingly diverse sets of modalities and intuitive forms of interacting with computing devices. Each of these frameworks looks at the human-computer interaction from a different angle, providing its unique advantages that do not necessarily contradict other frameworks but offer complementary possibilities and increase flexibility in designing new experiences. Despite this flexibility, these frameworks and the related interfaces generally offer modalities as separate and isolated features [7–9].

Humans use their senses and modalities in a holistic way, working together to achieve goals. The notion of wholes versus parts in different domains is not a new one. It has been around since Plato discussed the relationship between parts and the whole produced by them [10]. The term "holism" was introduced by Smuts in 1926 [11]. The idea is presented based on how the world works and how it consists of many important parts that have a tendency to result in wholes that are different from those parts. The holistic


M. Kurosu et al. (Eds.): HCII 2022, LNCS 13516, pp. 18–33, 2022.
https://doi.org/10.1007/978-3-031-17615-9_2

notion and whole-part differences play a key role in human experiences as we work with our modalities as parts of a whole experience.

The notion of holistic user experiences is not a new one either. Some researchers argue that user experiences are holistic by nature and that demands holistic criteria for their evaluation [12]. Even in the new technologies such as driverless cars, the concept of holistic user experience is being introduced [13]. However, HCI is bigger than just user experience, and this requires an adequately defined HCI framework that outlines the principles of holistic design and interaction. While there are many design guidelines and ideas for multimodal systems [1–3, 14], there is a lack of a unifying design methodology that is based on a holistic approach to user experience and interaction design.

In this paper, we start with a thorough literature review of HCI frameworks and the notions of multimodality and holism. We argue that existing approaches, while very helpful in certain applications, suffer from one or more of the following shortcomings:

1. They do not integrate various modalities within a given experience.
2. They do not provide a seamless transition between modalities within one task.
3. They do not offer modalities in a way that is consistent (performing the same tasks and achieving the same objective) and yet complementary (offering unique advantages).
4. They do not promote a design based on personalization.

To address these shortcomings, we propose Holistic Multimodal Interaction and Design (HMID), an HCI framework for designing holistic interactions. We initially conceived of HMID based on three principles of integration, seamless transition, and consistency. To investigate the guiding principles and effectiveness of the HMID framework, we conducted a user study that resulted in a fourth principle, personalization. The study focused on an imaginary retail experience as an example. We used illustrative storyboards [15] to simulate and present three retail scenarios. These scenarios follow a character who ends up purchasing a product. Each scenario is done through a separate approach: (1) The character engages separately with just two basic modalities in the retail experience. (2) The character engages in an isolated multimodal retail experience with a variety of modalities available. (3) The character engages in a holistic, multimodal retail experience (HMID). Our study incorporated a survey to see which scenario/approach was more successful with respect to the following goals:

1. Increasing sales
2. Increasing brand awareness
3. Increasing customer engagement

The results of our study suggested that a more multimodal experience using different technologies was perceived to increase sales, brand awareness, and customer engagement. However, there was not much difference between multimodal and HMID approaches. Upon reflection, we associated this with some participants not being prepared for or interested in a full HMID experience the way we had designed. As such, HMID should be flexible and adaptive, catering to the user's needs. This resulted in our fourth guiding principle, personalization.

The following sections review the relevant literature to provide a theoretical framework for HMID, then describe and discuss our study and its findings.

## 2 Related Work

### 2.1 Holism and Whole-Part Theories

The theory of part-whole originated as a purely philosophical concept in an attempt to understand the universe and human societies. The oldest trackable mention was in Plato's work, who discussed it informally [10]. However, the first formal discussion of the idea of parts and wholes relationships was in Aristotle's work when he established metaphysics [16]. He highlighted that the whole is not the same as its parts by highlighting that it does not necessarily result in a whole which is the sum of its parts [10]. In 1901, Husserl's third Logical Investigation was one of the first to transform the part-whole theory into a formal theory [10]. This theory was loosely referred to as the Part-whole theory before becoming a whole subfield of ontology commonly known as Mereology [17].

The original formulation of mereology as a field is credited to Stanislaw Leśniewski in 1916 [17]. However, since then, there have been many extensions of mereology that are referred to by different names [17]. By definition, mereology is the field that studies parthood relationships, the logical properties of the parts, and the whole they construct. The interrelationships between these parts and the whole they form are usually called mereological relationships. The thorough study of these relationships helped researchers produce more profound interpretations of these relationships, which helped in advancing many sciences, such as set theory in math [10]. This theory is the ground on which Object-Oriented Programming and the notion of classes are formed.

In 1926, Jan Smuts [11] introduced the term "holism." The notion is based on how nature works and how it consists of many important parts that have a tendency to form wholes. He argues in his philosophical theory that nature tends to produce wholes from the ordered grouping of the parts or units and that this tendency has been seen throughout existence. The hierarchical structure of entities (wholes vs. parts, sometimes referred to as "wholism") is related but not the same as the other aspect of this notion that wholes are more than the mere sums of the existing parts.

In psychology, the Gestalt theory emerged in the early 20$^{th}$ century, in contrast to elementalism and atomism, which tend to see all objects and ideas as a collection of elementary building blocks [18]. It promotes the notion of "whole is more than its parts." Gestalt principles of proximity, similarity, figure-ground, continuity, closure, and connection can be used to describe how humans perceive objects and environments.

Holistic and Gestalt principles offer insights on how to design multimodal systems in a holistic way. While connection and proximity are about proper integration, similarity, figure-ground, and continuity relate to how the parts should be consistent yet complementary.

### 2.2 Main HCI Frameworks

Early computers relied on text input/output and command lines for interacting with the user. The introduction of Graphical User Interfaces (GUI) was the basis of WIMP

(Windows-Icon-Menu-Pointer), the first major HCI framework [3]. The term WIMP was coined by Merzouga Wilberts in 1980, and the framework offered a new angle to look at computing devices; it compared computers to desktops as the most common workspace, instead of the old way of comparing them to typewriters. As a result, new forms of interaction were introduced that were symbolically related to real-life experiences (e.g., folders on desktop) but still relied on visuals and hand actions and did not quite reflect real-life interactions.

WIMP was followed by a series of other frameworks (generally referred to as Post-WIMP) that tried to move away from the 2D widget-based interactions to add new modalities [19]. Tangible User Interfaces (TUI) are based on the integration of tangible modalities and the notion of embodiment [4]. TUI's special angle was that human beings need to touch and feel things to experience them. This opened the door to the utilization of haptic technologies and expanded to offer new forms of wearable technologies. However, the focus of this framework, like the rest of the frameworks, was using modalities in combinations or isolation to achieve embodiment. This is an important approach that offers significant insights but does not necessarily provide the full picture.

Augmented/Virtual Reality (AR/VR) and 3DUI looked at HCI from the angle that the world is 3D and not 2D, and therefore, the interactions should move out of the 2D widget type of interactions to an extended form of reality. Milgram et al. [20] based this framework on the reality–virtuality continuum to include both virtual and augmented realities. These realities are based on adding visual 3D modalities into the environment. As a result, this framework offered some interesting new modalities and introduced new ways of training, gaming, and so on. AR/VR/3DUI do not necessarily provide new interaction modalities but a new environment to interact. They are commonly associated with head tracking, hand controllers, head-mounted displays, and various other forms of interaction. There is very limited research on guiding principles for combining modalities in AR/VR/3DUI.

Ubiquitous Computing, and related concepts such as Ambient Computing and the Internet of Things, influenced HCI significantly [5]. This framework is based on the idea that users are surrounded by an increasing number of smart devices, each one of these devices has a function it fulfills in everyday life, and by adding interconnectivity between these devices, regardless of the type of connection, it could offer new possibilities and experiences. As the internet became the most dominant and accessible means of communication, it led to the establishment of the IoT framework by looking at it from a different angle that incorporated the internet as a means of communication. This framework showcases the importance of looking at the changing functions of parts versus the whole they constitute. However, it was applied on a very narrow application, and this highlights the importance of looking at HCI from a holistic broader angle.

Adaptive User Interfaces (AUI) are based on using intelligent adapting mechanisms to overcome usability issues [21, 22]. The angle this framework looked at was the idea that human beings adapt to each other, and the environment based on their interactions and, as a result, interactions and communications become better. While this framework is focused on the customization of interactions, it would be very interesting to look at it holistically to assess if the interrelations between its modalities would affect this customization positively or negatively in different scenarios.

Natural User Interfaces (NUI) framework [6] looked at user interactions from the angle that HCI should feel natural by incorporating the user's natural interactions. Examples of these natural interactions are gaze, speech, touch, and gesture. Similarly, the Multimodal User Interfaces (MUI) framework [1, 2] is based on the simple idea that human beings do not experience the world through one modality. Therefore, it encouraged multimodality instead of focusing on one form of modalities which was common in 2D widgets. By emphasizing this angle, most experiences turned from visual to auditory to include new unique forms of modalities. While there are many studies within the realm of NUI and MUI, there is limited research on how to design multimodal experiences in a holistic way, i.e., how to relate modalities to each other. The focus in this area has been on special applications such as AR/VR or conversational agent [23, 24], or on multimodal integration [14], i.e., when to use multiple modalities, as opposed to what design principles to follow so that the modalities can offer a holistic experience.

## 3 Research Approach

### 3.1 A Framework for Designing Holistic Multimodal Interaction

As discussed earlier, while there are many HCI frameworks, the majority are focused on the presence of modalities or the best way to utilize them in specific scenarios. Additionally, they tend to examine these modalities in isolation to solve different sets of problems. The effect of different modalities as a whole to solve a problem is different from the effect of each in isolation. Furthermore, there have been many new trends that are becoming common in new experiences. However, there is no integrated framework that systematically examines these trends and notions thoroughly as parts of a whole, rather than being optional afterthoughts.

The notion of holistic frameworks is not a new one. However, it has been mainly focused on user experience, which is only one element of HCI. These user experiences do not consider functionality or the interrelation and dynamics between modalities or mereological parts as part of the design process. For example, Pallot et al. [12] investigated different attributes of user experience design in living labs. Living labs are labs designed for studying user experiences. As a result, they produced a sample experience-illustration of a limited number of factors in a holistic user experience. They came up with a list of 10 experience types, 22 elements, and about 80 properties. They identified the following experience types: perceptual, cognitive, reciprocal, social, emotional, cultural, empathetical, technological, economical, legal, and ethical. Then each one of these experience types had elements and properties. However, this and similar studies focus only on understanding the user experience, and they do not offer guidelines on how to design multimodal experiences.

Inspired by the whole-part theories and the notions of holism and Gestalt, we propose the Holistic Multimodal Interaction and Design (HMID) framework. HMID borrows the principles of holism and holistic solutions in other fields and defines three initial principles for holistic interaction design in HCI:

1. **Integration**: This principle is the foundation of multimodal systems and states that a multimodal experience should bring together different modalities as one package.

It means that the multimodal experience is not a set of separate actions, each with one modality, but a single design that uses those modalities where and how they are more suitable.

2. **Seamless transition**: As a follow-up to integration, seamless transition means that the user should be able to switch from one modality to another at any time, even within a task, without much effort or loss of data.
3. **Consistency**: The third principle combines two seemingly opposite aspects, being compatible with being complementary. It means that while different modalities follow the same objective and communicate the same message, they complement each other by providing different advantages.

We performed an initial evaluation of HMID and its guiding principles (compared to limited modality and isolated multimodal cases) through a simulated retail experience using storyboards, as described in Sect. 4. The findings of this study (Sect. 5) suggested a fourth principle, personalization, which will be discussed in Sect. 6. Before providing the details of the study, we will briefly review the concept of storyboards and the initial validation process that prepared us for the main study.

### 3.2 Using Storyboards

To investigate the potential of a holistic approach in the context of a retail experience, we followed the approach of illustrating hypothetical scenarios by using storyboards. In the absence of actual system implementation, due to logistic restrictions and the COVID pandemic, storyboards offer a simulated experience for users to interact. We selected to use storyboards to explain what a user would experience in a common visual language so that participants can easily understand the interactions with the system. Figure 1 shows the first three panels of a simulated experience.


**Fig. 1.** Sample storyboard.

The design process of the storyboards followed the guidelines established by Truong et al. [15] in that we used short textual descriptions to describe the process, included people to explain the interactive experience, indicated the passage of time only when necessary, and used the minimum level of detail required to understand the scenario. All created storyboards are freely available online:
https://www.dropbox.com/sh/9rzfp4d71t2tsj0/AADdehzgIB2pzqR_xuvKoxKca.

In each of the three storyboard case scenarios (limited modality, isolated multimodal, and HMID), there is a customer engaged in a retail fashion shopping experience. The user goes through several situations that eventually lead to a product sale.

To ensure that participants understood the storyboards, we conducted a pre-study survey with five participants (3 males and 2 females) ranging in age from 30 to 39 years old. A questionnaire depicting early sketches of the three scenarios was sent asking for feedback on the clarity of the descriptions and the understanding of the scenarios. Some examples of the questions asked were: "Does the scenario make sense?", "Are the user's actions clear?", and "Do you have any suggestions for improving the pictures?". Participants were also invited to provide any general comments they had for improvement. To summarize, below were the comments we received and considered while iterating on the storyboard designs:

- Transition across the devices using voice commands is seamless and intuitive.
- The treasure hunt idea was well-liked and stimulated the most interest, yet some could not relate it to the retail experience.
- Overall, the pictures were clear. But because there is so much going on, it can be hard to keep track of. As such, in the final design, some panels were combined to shorten the story while still retaining the main idea.
- Many found that the use of QR codes was a good idea for encouraging customer engagement and interaction with the product.
- The fact that there were different payment options was not clear.

## 4 Main Study

We conducted an online survey to evaluate the preferences for scenarios. The independent variables are the three case scenarios (1) two modalities, (2) isolated multimodal, and (3) holistic multimodal. The dependent variables are increased sales, increased brand awareness, and increased customer engagement.

### 4.1 Participants

A total of 140 participants completed the survey. However, 21 of the responses were discarded as the descriptions did not properly describe the scenario, answers were too brief (e.g., one-word answers), the completion time was too short (e.g., under 3 min), or attention check questions were missed. The final data set contained 119 responses (75 men and 44 women) ranging in age from 18 to 50 years old ($M = 34.41$ years, $SD = 7.65$ years). 108 participants reported they enjoyed shopping, while 11 did not. Among the 108 respondents, 41 of them prefer to shop alone, 18 of them prefer to shop with friends and 49 prefer to shop both with friends and alone.

### 4.2 Materials and Instruments

A retail experience was the subject of our study. The goal was to design a marketing and sales campaign to increase sales, customer engagement, and brand awareness. A

variety of modalities were included, such as printed material, website, mobile apps, AR advertising, a multi-platform treasure hunt to collect points, and in-store interaction.

The storyboards show a character and his interactions with technologies that were corresponded to each of the three case scenarios:

1. For the "two modalities" scenario, the user engages separately with two modalities of the retail experience.
2. For "isolated multimodal", the user engages in an isolated multimodal retail experience.
3. For "holistic multimodal", the user engages in a holistic, multimodal experience. This case is designed using the three principles of integration, seamless transition, and consistency.

The principles of HMID were implemented in the storyboard interactions by showing how the user can seamlessly switch between the different kinds of interaction modalities at hand. For HMID, we showed the same experience being available using all modalities, but each with its own conveniences. For example, a desktop browser could be simply used to collect points by answering questions. But an AR-based mobile app, while more complex, could be more convenient and allow a richer experience. We used storyboards to simulate the user experience. Figure 2 shows examples of three HMID principles at work.

To compare the performance of HMID compared to two other approaches, we considered three evaluation criteria: sales, customer engagement, and brand awareness (Table 1). Following each scenario, participants indicated their level of agreement on a 7-point (1 = strongly disagree to 7 = strongly agree) Likert scale.

### 4.3 Recruitment and Procedures

After receiving clearance from our institutional ethics review committee, the recruitment for participants commenced. Participants were recruited using Amazon Mechanical Turk (MTurk) [25]. MTurk is an online crowdsourcing platform where "workers" complete small tasks for monetary reward [26]. It is a popular method for recruiting large and diverse participants in many HCI studies [27–29]. Studies show that most workers participate out of interest or to pass the time, rather than for the sake of the reward, making these participants a good source for collecting data [30]. Participants had to be over 18 years old, must understand English, and own a computer or a mobile device with an internet connection to complete the questionnaire. A link to the questionnaire was hosted online (Qualtrics, Provo, UT) and takes approximately 30 min to complete. The questionnaire consisted of three parts. Part I collected demographic-related information (e.g., age, gender, and level of education), as well as information about shopping behaviours and preferences. Part II provided some background context to explain the scenarios, followed by the storyboards where participants were invited to look at the drawings and indicate their level of agreement or disagreement on a 7-point Likert scale to questionnaire statements that evaluated the dependent variables: increased sales, brand awareness, and customer engagement. Participants were also asked an open-ended question to share any further comments about the experience after each case scenario.


**Fig. 2.** HMID principles implemented in storyboards. Panels 7–10 and 25–28 show integration, seamless transition, and consistency.

**Table 1.** Questionnaire statements that evaluated dependent variables.

| Dependent variable | Questionnaire statements |
|---|---|
| Increase sales | I felt that the technological features were helpful in promoting increased sales |
| | I felt that the product promotion and the use of technology helped increase the sales |
| Increase customer engagement | I felt engaged with the brand campaign |
| | I felt delighted by the brand's promotional experience |
| | I felt that the technology features were easy for me to engage with the brand |
| Increase brand awareness | I felt the brand's promotional campaign and the use of technology did well in increasing brand awareness |

To ensure that respondents understood each case scenario portrayed in the storyboards, in addition to asking them to read the text provided underneath each storyboard, they were also asked to write a short description of what they thought was happening in the images before they provided a rating. At the end of each scenario, respondents were invited to provide any other thoughts they had. Attention check questions adapted from the Conscientious Responder Scale [31] were integrated into the questionnaire to verify that respondents were reading and answering the questions carefully. Participants received $2.00 USD as compensation for participating in the study.

### 4.4 Statistical Analysis

Analyses were performed using IBM SPSS for Windows Version 27.0 (SPSS, Chicago, Illinois), and prior to analyses, including the filtering process described in 4.1, the data were screened for missing values, outliers, and out-of-range values. Descriptive statistics and the Shapiro-Wilk's test of normality [32] were calculated for each of the three retail experience cases (two modalities, isolated multimodal, and holistic multimodal) with respect to the three dependent variables (increase sales, increase brand awareness, and increase engagement).

## 5 Results

### 5.1 Quantitative Results

Descriptive statistics and the Shapiro-Wilk's test were conducted to examine the normality of the sample (Table 2). Based on the shape of the distributions and the results of the Shapiro-Wilk's test, non-parametric tests were selected for analysis. An alpha level of 0.05 was set for all statistical tests.

To explore the data, 100% stacked bar charts were generated to compare each of the case scenarios with respect to the dependent measures (Fig. 3). This provided a visual representation to examine the overall distribution. Based on Fig. 3, we can see that participants preferred a more holistic experience for each of our dependent measures. It is interesting to note that for each dependent measure, the percentage between Case 2 and Case 3 showed minor differences.

A Friedman test was conducted to examine the differences between each of the three cases with respect to our dependent measures. Results showed that there were statistically significant differences between the cases for increased sales and customer engagement, but there were no statistically significant differences for brand awareness (Table 3).

Next, post-hoc tests, using the Wilcoxon signed-rank tests, were conducted between each pair of cases separately to further examine where the differences occurred.

Table 4 summarizes the results. For increased sales, case 2 is better than 1, likewise, case 3 is also better than 1, but no difference between case 2 and 3. This suggests that a holistic experience will likely increase sales. For brand awareness, case 2 is better than 1, but no differences between case 1 vs. 3 and 3 vs. 2. This suggests that a holistic experience may be similar to other cases for increasing brand awareness. For customer engagement, there were no differences between case 1 and 2, but there was a statistically significant difference between case 1 vs. 3, and 2 vs. 3. This suggests that offering a holistic experience might not necessarily increase the level of engagement.

**Table 2.** Descriptive statistics for the level of agreement for the three dependent measures: increased sales (IS), brand awareness (BA), and customer engagement (CA), with respect to each of the three cases: two modalities (TM), isolated multimodal (IM), and holistic multimodal (HM), and the Shapiro-Wilk's test of Normality.

| $N = 119$ | | Descriptive Statistics | | | | | Shapiro-Wilk's test | | |
|---|---|---|---|---|---|---|---|---|---|
| Case | Measure | M | SD | Mdn | Skewness | Kurtosis | W | df | p |
| TM | IS | 10.60 | 2.73 | 11.00 | − .69 | − .06 | .93 | 119 | .00 |
| | BA | 5.43 | 1.54 | 6.00 | − .96 | .29 | .86 | 119 | .00 |
| | CE | 15.00 | 3.85 | 15.00 | − .52 | .32 | .96 | 119 | .00 |
| IM | IS | 11.27 | 2.41 | 12.00 | − .80 | .16 | .91 | 119 | .00 |
| | BA | 5.79 | 1.21 | 6.00 | − .97 | .56 | .85 | 119 | .00 |
| | CE | 15.47 | 3.64 | 16.00 | − .65 | .18 | .95 | 119 | .00 |
| HM | IS | 11.41 | 1.88 | 12.00 | − .39 | − .52 | .94 | 119 | .00 |
| | BA | 5.63 | 1.34 | 6.00 | − 1.06 | .91 | .86 | 119 | .00 |
| | CE | 11.74 | 3.80 | 11.00 | .67 | − .04 | .94 | 119 | .00 |

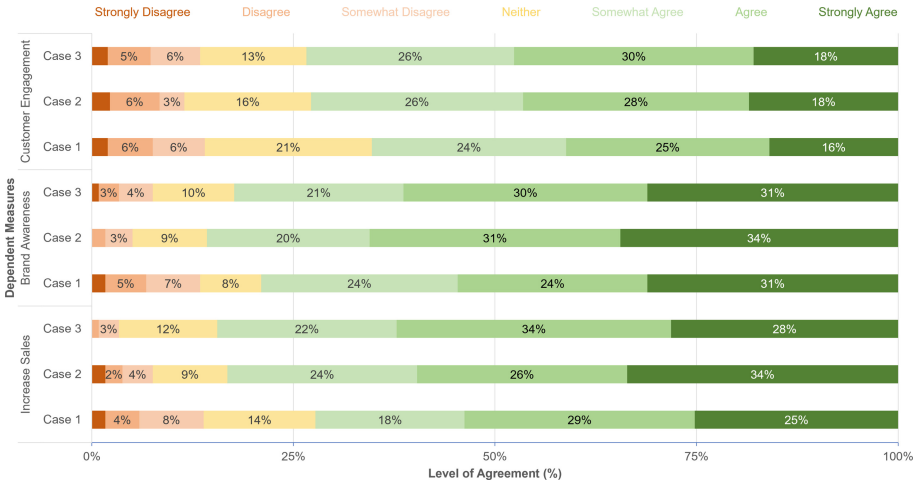


**Fig. 3.** 100% stacked bar charts for the level of agreement on dependent measures with respect to three case scenarios: (1) two modalities, (2) isolated multimodal, and (3) holistic multimodal ($N = 119$).

## 5.2 Qualitative Results

At the end of each case scenario, participants were given the opportunity to provide their feedback. This was helpful as it gave us insight into how the increased use of technologies, that is, from a single modality to an HMID experience, would be received from this population sample. We conducted open coding, the initial interpretive process by which

**Table 3.** Mean rank results of a Friedman test for each of the three cases with respect to the dependent measures.

| | Mean Ranks | | | Friedman test statistics | | |
|---|---|---|---|---|---|---|
| Measures | Case 1 | Case 2 | Case 3 | $\chi^2$ | df | p |
| Increases Sales | 1.81 | 2.13 | 2.05 | 10.15 | 2 | .01 |
| Brand Awareness | 1.91 | 2.13 | 1.96 | 4.870 | 2 | .09 |
| Customer Engagement | 2.24 | 2.37 | 1.40 | 71.29 | 2 | .00 |

**Table 4.** Mean rank results of Wilcoxon's signed-rank tests for each of the three cases: (1) two modalities, (2) isolated multimodal, and (3) holistic multimodal, with respect to the dependent measures: increase sales, brand awareness, and customer engagement.

| | | Mean Ranks | | Wilcoxon test | |
|---|---|---|---|---|---|
| Measures | Case comparisons | Positive | Negative | $\chi^2$ | p |
| Increases sales | 1 vs 2 | 37.14 | 36.79 | − 2.61 | .01 |
| | 1 vs 3 | 40.71 | 34.71 | − 2.56 | .01 |
| | 2 vs 3 | 44.60 | 32.52 | − .49 | .62 |
| Brand awareness | 1 vs 2 | 34.96 | 27.19 | − 2.52 | .01 |
| | 1 vs 3 | 37.43 | 33.20 | − 1.08 | .28 |
| | 2 vs 3 | 36.70 | 35.55 | − 1.50 | .13 |
| Customer engagement | 1 vs 2 | 50.41 | 45.20 | − 1.09 | .28 |
| | 1 vs 3 | 31.65 | 62.72 | − 6.92 | .00 |
| | 2 vs 3 | 28.26 | 62.62 | − 7.92 | .00 |

raw research data are first systematically analyzed and categorized [33], followed by axial coding, a qualitative research technique that involved relating data together to uncover codes, categories, and sub-categories grounded within participants' voices within the collected data [34]. This method was appropriate not only to organize the data but was also useful for deriving potential new theories and concepts, insights that may be of value. With open coding, we deduced three categories: (1) Positives, (2) Negatives, and (3) Concerns. "Positives" relate to the positive attitudes towards the idea and user experience, while "negatives" describe the doubts and skepticisms about the idea and user experience. "Concerns" are worries that users had about the experience.

Next, we further compartmentalized participants' responses using axial coding. This was achieved by recognizing like-patterns in participants' answers, thereby collecting and categorizing them appropriately. Axial coding revealed Fun/Enjoyment, Engagement, and Immersive under Positives, Too Complex, Overwhelming, Waste of Time and Annoyed under Negatives, and Privacy and Tracking under Concerns.

In general, we found that most participants who provided feedback related to "fun and engagement" stated that they welcomed the increase in technology. For example, in Case 2, one participant (P1) wrote: "*I thought the technology made shopping a more pleasant experience in this scenario*" when comparing it with Case 1. In Case 3, another participant (P17) said: "*I thought there was a lot of technology in this scenario, and it all worked together to provide a better shopping experience*". Furthermore, participants generally thought that the experience simulated in Case 3 was engaging and exciting if the process was efficient. For example, one participant (P90) said that "*as long as steps and rewards are easy to follow and understand, then I love utilizing all the platforms to buy items*". Another participant (P33) thought that "*the brand did a great job engaging a certain segment of its customer base into buying more and being more excited by its use of technology as a marketing technique*". One more participant (P7) wrote: "*I like the idea of full use of technology like the last one and making it seamless between devices*". This suggests that the holistic experience portrayed in Case 3 was well-received.

Participants also reported negative experiences. Participants expressed that the excessiveness and the steps to use the technology to achieve the goal of purchasing products would be overwhelming. For example, in Case 2, one participant (P10) said that there are "*too many steps and too complicated*". Another participant (P38) explained that it "*seems like a lot of needless extra work to buy item*". For design, these comments suggest that it is important to make the process as simple as possible.

Interestingly, there was only one participant (P41) who raised concerns about data privacy, saying that: "*I have a concern with tracking and privacy*".

## 6 Discussion and Redesign

The result from the data supports the idea that a more multimodal experience using different technologies favour increased sales, brand awareness, and customer engagement. However, there is not much difference between multimodal Case 2 and my hypothesized HMID Case 3. There could be a variety of reasons for this. Let me explain.

It could be that the storyboards themselves needed to convey a better narrative of HMID. But a more likely explanation is that not all participants were comfortable with using various modalities/technologies in the HMID case. Reflecting on this possibility, we hypothesized that a holistic experience should be flexible and adaptive, catering the user's needs.

To informally verify this hypothesis, we ran a small mini-study with only three participants. We picked four sections of the HMID case where the user was using different modalities, and offered options such as a more convenient technology or a more familiar one. For example, Fig. 4 shows a section of the storyboard where the user received a notification on his smart watch that there is an AR-based point to collect as part of the treasure hunt. We asked if the existence of an option to save this location for the future, instead of having to use AR, adds flexibility, and so improves the HMID experience. All participants agreed with comments such as "Yes, it would improve the HMID scenario because any sort of flexibility or convenience provides a better experience for the user" (P2).

The findings of the main and follow-up studies encouraged us to add personalization as the fourth principle of holistic, multimodal interaction and design.

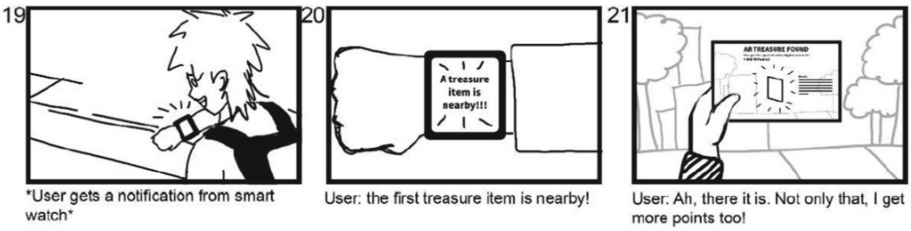


**Fig. 4.** Storyboard Sample used in follow-up mini-study.

## 7 Conclusion

Based on a review of literature on HCI frameworks and whole-part theories, we proposed a holistic, multimodal interaction and design framework based on the principles of integration, seamless transition, and consistency. Using a conceptual storyboard approach and survey research methodology, we compared the perceived effect of this framework with two other cases (limited modality and isolated multimodal) in a simulated retail scenario with the objectives of increasing sales, brand awareness, engagement, and overall satisfaction. The results from both quantitative and qualitative analyses suggested that multimodal and holistic experiences are more effective with regard to the objectives. Further reflection suggested that personalization should be added as a fourth principle to allow flexibility in holistic experiences.

As society increasingly adopts emerging technologies (such as wearables, AR/VR, amongst others), a holistic and cohesive approach for the seamless integration of data from one multimodal to another is necessary. Although this research study was specifically catered to the retail shopping scenario, we believe that HMID can potentially lay the foundation for new HCI paradigms for IoTs. Plausible future HMID research directions include education, sports, and how we engage and interact with information media platforms.

## References

1. Nigay, L., Coutaz, J.: A design space for multimodal systems. In: Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '93. ACM Press, New York, New York, USA, pp 172–178 (1993)
2. Oviatt, S.: Multimodal Interfaces. human-computer Interact Handb Fundam Evol Technol Emerg Appl **14**, 405–430 (2012). https://doi.org/10.1201/b11963-22
3. Kruschitz, C., Hitz, M.: Human-computer interaction design patterns: structure, methods, and tools. Int J Adv Softw **3**, 225–237 (2010)
4. Ishii, H., Ullmer, B.: Emerging frameworks for tangible user interfaces. IBM Syst J **39**, 915–931 (2000)
5. Weiser, M.: The computer for the 21st century. Sci Am **265**, 94–104 (1991). https://doi.org/10.1038/scientificamerican0991-94
6. Jain, J., Lund, A., Wixon, D.: The future of natural user interfaces. Conf Hum Factors Comput Syst - Proc 211–214 (2011). https://doi.org/10.1145/1979742.1979527

7. Gorlewicz, J.L., Tennison, J.L., Uesbeck, P.M., et al.: Design guidelines and recommendations for multimodal, touchscreen-based graphics. ACM Trans. Access Comput. **13** (2020). https://doi.org/10.1145/3403933
8. Reeves, L.M., Martin, J.C., McTear, M., et al.: Guidelines for multimodal user interface design. Commun ACM **47**, 57–59 (2004)
9. Sarter, N.B.: Multimodal information presentation: design guidance and research challenges. Int J Ind Ergon **36**, 439–445 (2006). https://doi.org/10.1016/j.ergon.2006.01.007
10. Snyder, L.J.: Stanford Encyclopedia of Philosophy Stanford Encyclopedia of Philosophy: Implicit Bias. 1–22 (2017)
11. Smuts, J.C.: Holism and Evolution. Macmillan (1926)
12. Pallot, M., Pawar, K.: A holistic model of user experience for living lab experiential design. 2012 18th Int Conf Eng Technol Innov ICE 2012 - Conf Proc. (2012) https://doi.org/10.1109/ICE.2012.6297648
13. Luxoft Developed a Holistic Automotive User Experience Solution Based on SmartDeviceLink | Business Wire. https://www.businesswire.com/news/home/20131111006225/en/Luxoft-Developed-a-Holistic-Automotive-User-Experience-Solution-Based-on-SmartDeviceLink. Accessed 9 Jun 2022
14. Turk, M.: Multimodal interaction: a review. Pattern Recognit Lett **36**, 189–195 (2014). https://doi.org/10.1016/j.patrec.2013.07.003
15. Truong, K.N., Hayes, G.R., Abowd, G.D.: Storyboarding: an empirical determination of best practices and effective guidelines. Proc. Conf. Des. Interact Syst. Process. Pract. Methods, Tech. DIS **2006**, 12–21 (2006)
16. Mayhew, R.: Part and whole in aristotle's political philosophy. J Ethics **1**, 325–340 (1997). https://doi.org/10.1023/A:1009743012461
17. Heller, M., Simons, P.: Parts: a study in ontology. Philos. Rev. **100**, 488 (1991). https://doi.org/10.2307/2185078
18. Mather, G.: Foundations of perception. 404–427 (2006)
19. Van Dam, A.: Post-WIMP user interfaces. Commun ACM **40**, 63–67 (1997). https://doi.org/10.1145/253671.253708
20. Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Mixed Reality (MR) Reality-Virtuality (RV) Continuum. Syst Res 2351, 282–292 (1994)
21. Raheel, S.: Improving the user experience using an intelligent Adaptive User Interface in mobile applications. 2016 IEEE Int Multidiscip Conf Eng Technol IMCET 2016 64–68 (2016). https://doi.org/10.1109/IMCET.2016.7777428
22. Hui, K.K., Liu, J., Wong, C.K.: An adaptive user interface based on personalized learning. IEEE Intell. Syst. **18**, 52–57 (2003)
23. Krishnaswamy, N., Pustejovsky, J.: An evaluation framework for multimodal interaction. Lr 2018 - 11th Int Conf Lang Resour Eval 2127–2134 (2019)
24. Rakkolainen, I., Farooq, A., Kangas, J., et al.: Technologies for multimodal interaction in extended reality—a scoping review. Multimodal Technol. Interact **5** (2021). https://doi.org/10.3390/mti5120081
25. Keith, M.G., Tay, L., Harms, P.D.: Systems perspective of amazon mechanical turk for organizational research: review and recommendations. Front Psychol **8** (2017). https://doi.org/10.3389/fpsyg.2017.01359
26. Difallah, D.E., Catasta, M., Demartini, G., et al.: The dynamics of micro-Task crowdsourcing: The case of amazon MTurk. WWW 2015 - Proc 24th Int Conf World Wide Web 238–247 (2015). https://doi.org/10.1145/2736277.2741685
27. Orji, R., Tondello, G.F., Nacke, L.E.: Personalizing persuasive strategies in gameful systems to gamification user types. Conf Hum Factors Comput Syst - Proc 2018-April (2018). https://doi.org/10.1145/3173574.3174009

28. Hirsh, J.B., Kang, S.K., Bodenhausen, G.V.: Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. Psychol. Sci. **23**, 578–581 (2012). https://doi.org/10.1177/0956797611436349
29. Jia, Y., Xu, B., Karanam, Y., Voida, S.: Personality-targeted gamification: a survey study on personality traits and motivational affordances. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16. ACM Press, New York, New York, USA, pp 2001–2013 (2016)
30. Buhrmester, M., Kwang, T., Gosling, S.D.: Amazon's mechanical turk: a new source of inexpensive, yet high-quality data? Methodol issues Strateg Clin Res (4th ed) 133–139 (2015). https://doi.org/10.1037/14805-009
31. Marjanovic, Z., Struthers, C.W., Cribbie, R., Greenglass, E.R.: The conscientious responders scale: a new tool for discriminating between conscientious and random responders. SAGE Open **4** (2014). https://doi.org/10.1177/2158244014545964
32. Mohd, N., Bee, Y.: Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J. Stat. Model Anal. **2**, 13–14 (2011)
33. Mills, A., Durepos, G., Wiebe, E.: Coding: open coding. In: Encyclopedia of Case Study Research. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States (2013)
34. Allen, M.: Axial coding. In: The SAGE Encyclopedia of Communication Research Methods. SAGE Publications, Inc, 2455 Teller Road, Thousand Oaks California 91320 (2017)