



CATCHWORD

Multiexperience

Ulrich Gnewuch · Marcel Ruoff · Christian Peukert · Alexander Maedche

Received: 9 July 2021 / Accepted: 27 June 2022
© The Author(s) 2022

Keywords Multiexperience · Device · Modality · Multi-device · Multimodality · Human–computer interaction

1 Introduction

The development of new devices, such as smart speakers or wearables, and recent advances in artificial intelligence (AI) that facilitate more natural interactions via speech or gestures are changing the interplay between user, task, and technology within information systems (IS). Today, people own multiple different devices, such as personal computers, smartphones, tablets, or smart speakers, and use them interchangeably, often switching between multiple devices in order to complete a task (Levin 2014; Westcott et al. 2020). In addition, many of these devices afford users new ways of interacting with them through touch, speech, or gestures (Turk 2014). For example, customers can shop at Amazon using multiple devices and multiple interaction modalities as well as combinations thereof (e.g., Amazon’s Echo Show devices combine speech interaction with a touchscreen display). The same trend can be found at the workplace, where employees can, for example, use natural language to interact with enterprise resource planning (ERP) systems (e.g., SAP CoPilot) or business intelligence and analytics (BI&A) systems (e.g., Tableau Ask Data).

This trend has not gone unnoticed by market research firms that recently introduced the term *multiexperience* (MUX) as a key area of strategic importance in the next years (Gartner 2019). There also is a plethora of existing research related to MUX that provides insights into how users interact with IS across different devices and modalities and how to design for MUX (Brudy et al. 2019; Li and Zhang 2005; Turk 2014; Zhang et al. 2009). However, what is new today is that the sheer number of available devices and mature modalities presents an opportunity – and challenge – to better meet users’ needs and preferences when they interact with an IS to perform tasks. Similar tasks can be performed quite differently and may result in different outcomes, depending on the devices used (e.g., smartphones vs. smart speaker) and modalities available (e.g., clicking on a screen vs. speech input) (Diederich et al. 2020; Rzepka et al. 2020a). Thus, there is a need to improve our understanding of the nature of MUX and the roles of devices and modalities within IS. Considering that many devices (e.g., virtual reality headsets) and modalities (e.g., speech interaction) are now beginning to reach a level of maturity which allows widespread application, we believe the time is ripe to (re)define the concept of MUX.

The goal of this catchword is to build a bridge between the new term of MUX and existing research in our field in order to provide a solid conceptual grounding for MUX and identify future research opportunities for the BISE community. Drawing on the rich body of research on multi-device and multimodal IS in the fields of IS and human–computer interaction (HCI), we propose a clear conceptualization of MUX and provide a framework of three guiding paths toward MUX that may be equally useful for researchers and practitioners in the BISE community. Additionally, we describe several real-world examples to explain the benefits and challenges of each path in our

Accepted after two revisions by Christine Legner.

U. Gnewuch (✉) · M. Ruoff · C. Peukert · A. Maedche
Institute of Information Systems and Marketing, Karlsruhe
Institute of Technology, Kaiserstraße 89-93, 76133 Karlsruhe,
Germany
e-mail: ulrich.gnewuch@kit.edu

framework and offer practical guidance for moving along these paths toward MUX. Finally, we outline promising areas for future research and highlight the unique position of the BIASE community in capitalizing on these opportunities.

2 Conceptual Foundations: Multi-Device and Multimodal Information Systems

In the early days of personal computing, researchers and practitioners primarily focused their efforts on a single device – the personal computer (PC). However, already in 1991, Mark Weiser shared his vision of a world in which people can interact with content across multiple computing devices in different shapes and sizes (Weiser 1991). This idea served as inspiration for research that went beyond a single user at a single computer (Brudy et al. 2019). A classic example is Rekimoto’s seminal work from the late 1990s on interaction techniques that crossed device boundaries of multiple portable computers and displays on table/wall surfaces (Rekimoto 1997; Rekimoto and Saitoh 1999). Around the same time, commercial products such as the first BlackBerry device were introduced that offered new ways to do everyday tasks that would typically be done on a computer in an office (Lyytinen and Yoo 2002). For example, people could not only send and receive emails but also synchronize their schedule, tasks, and contacts with their PC – something that is commonplace nowadays but represented a major leap forward at that time. Since the advent of mobile devices in the 1990s in general, and smartphones in the 2000s in particular, the number and diversity of devices has grown significantly. In addition, many companies allow employees to use private devices for work purposes, and vice versa (Köffer et al. 2015). Today, the most popular devices range from PCs, smartphones, tablets, and TVs to smart speakers, smart-watches and augmented (AR) or virtual reality (VR) devices (GlobalWebIndex 2020). Each device is characterized by certain display capabilities (e.g., screen size), processing power, input/output modalities, and sensors (Levin 2014). Furthermore, people use devices much differently than they did 10 or 20 years ago. Since many tasks span multiple devices (Dearman and Pierce 2008), people use several devices simultaneously and switch between them to complete a single task (Brudy et al. 2019). For example, Netflix’s “Continue Watching” feature allows customers to start watching a movie on one device (e.g., a TV in the living room) and continue watching on another device (e.g., a smartphone or tablet) while commuting or waiting for an appointment (Netflix 2013). However, there are also several technical challenges associated with multi-device IS, particularly when it comes to sharing

information and keeping it consistent across multiple devices (Dong et al. 2016). Commercial solutions, for example, are often limited to devices within a particular manufacturer’s ecosystem (e.g., Apple) and there are only few open standards that support the integration of multiple devices (Brudy et al. 2019). Nonetheless, current trends suggest that researchers and practitioners cannot consider the PC, smartphone, or any other device as a standalone platform anymore, but need to understand and design for use patterns across multiple devices (Levin 2014).

An important distinguishing feature of the aforementioned devices is that they offer users a wide variety and diversity of interaction modalities (hereafter referred to as *modalities* for simplicity). Modality broadly refers to the type of communication channel used to convey or acquire information (Nigay and Coutaz 1993). This includes both input modalities (i.e., users providing data to the system) and output modalities (i.e., users receiving data from the system). Users provide input using their effectors (e.g., limbs, eyes, vocal system, head) and perceive output through their five senses (i.e., sight, hearing, touch, smell, and taste). For example, in the interaction with a PC, users primarily provide input through typing on a keyboard or mouse clicks using their fingers (i.e., limbs). Further input modalities, such as speech, mid-air gestures, and eye gaze, that leverage other effectors are possible but less common today. In terms of output modalities, users primarily rely on their visual sense since most applications on a PC feature a graphical user interface. Secondary output modalities that leverage other senses, such as audio output (e.g., “beeping” when an error occurs), may also play a role. Table 1 provides an overview of different categories of devices and their input as well as output modalities.

Although humans employ multiple senses and effectors to interact with the world around them, research in the fields of IS and HCI has historically been focused on unimodal interaction (i.e., using only a single input and a single output modality) (Liu et al. 2019; Turk 2014). An example is the PC that displays text on a screen with a keyboard for input. However, advances within the AI subfields of natural language processing and computer vision as well as affordable sensor technology are paving the way toward more natural interactions that replace or complement traditional modalities (Turk 2014). Multi-modality refers to the use of more than one input and/or output modality in the interaction (Nigay and Coutaz 1993). These modalities can be used simultaneously or sequentially during the interaction. For example, in the interaction with an Amazon Echo Show device, users can provide input via speech (e.g., “Alexa, what’s the weather like in Berlin today?”) and its touch screen (e.g., touching a button) and receive spoken output (e.g., “The current weather in Berlin is ...”) and visual output on the screen

Table 1 Examples of devices and input/output modalities

Devices	Input modalities						Output modalities ^a		
	Typing (e.g., keyboard)	Pointing (e.g., mouse, pen)	Touch	Speech	Mid-air gestures	Gaze	Vision	Audio	Haptic
Personal computer (desktop/laptop)	X	X	x	x	x	x	X	x	
Smartphone and tablet	x	x	X	x	x		X	x	x
Large interactive screens (e.g., Surface Hub)	x	x	X	x			X	x	
Smart speaker (e.g., Amazon Echo)				X				X	
Smartwatch (e.g., Apple Watch)			X	x			X		x
Augmented reality smart glasses (e.g., HoloLens 2)				x	X	x	X	x	
Virtual reality headset (e.g., Oculus Quest 2)		X		x	X		X	x	x

Bold entries are the prominent modalities of a device. ^aSmell and taste only exist in research prototypes today (Obriest et al. 2016)

Table 2 Overview of literature streams related to multiexperience

Research stream	Description	Important papers	Exemplary artifacts
Multi-Device Information Systems	Information systems that offer users the ability to use more than one device in the interaction. The use can be sequential or in parallel.	<ul style="list-style-type: none"> •Weiser (1991) •Rekimoto and Saitoh (1999) •Lyytinen and Yoo (2002) •Dearman and Pierce (2008) •Köffer et al. (2015) •Dong et al. (2016) •Brudy et al. (2019) 	<ul style="list-style-type: none"> •“Pick-and-Drop” (Rekimoto 1997) •BlackBerry email service (1999) •Play-along apps for TV shows (e.g., The Million Pound Drop; 2012) •Netflix’s “Continue Watching” feature (2013) •Apple’s “Continuity” feature on iOS and macOS devices (2014)
Multimodal Information Systems	Information systems that offer users the ability to use more than one input and/or output modality in the interaction. The use can be sequential or in parallel.	<ul style="list-style-type: none"> •Nigay and Coutaz (1993) •Oviatt (1999) •Lee et al. (2001) •Reeves et al. (2004) •Suh and Lee (2005) •Turk (2014) •Liu et al. (2019) 	<ul style="list-style-type: none"> •“Put That There” (Bolt 1980) •PalmPilot PDA (1997) •iPhone (2007) Smartwatches (e.g., Fitbit; 2015) •Amazon Echo Show (2017) •Microsoft HoloLens 2 (2019)

(e.g., a weather forecast). The key assumption is that “well-designed multimodal systems integrate complementary modalities to yield a highly synergistic blend in which the strengths of each mode are capitalized upon and used to overcome weaknesses in the other” (Oviatt 1999, p. 74). Indeed, research has shown that multimodality can lead to higher task performance (Lee et al. 2001) and improve learning outcomes (Suh and Lee 2005). Similar to multiple devices, the integration of input from multiple modalities is a key technical challenge (Reeves et al. 2004). The main reason is that due to the unique characteristics of each modality, there are no obvious points of similarity and therefore no straightforward ways to connect them (Turk 2014). Over the years, several so-called fusion approaches have been developed to address this challenge (Jaimes and Sebe 2007). In general, fusion can be performed at different levels, ranging from feature level (i.e., integrating

input signals) to higher semantic levels (i.e., integrating common meaning representations derived from different modalities) (for an overview, see Jaimes and Sebe 2007). Although significant progress has been made, technical challenges related to the integration of modalities remain and the development of fusion approaches continues to be an active area of research.

To summarize the conceptual foundations of MUX, Table 2 provides an overview of the two research streams on multi-device and multimodal IS with key papers and exemplary artifacts.

3 Three Paths Toward Multiexperience

Against the backdrop of the conceptual foundations, we define MUX as the user’s perceptions and responses

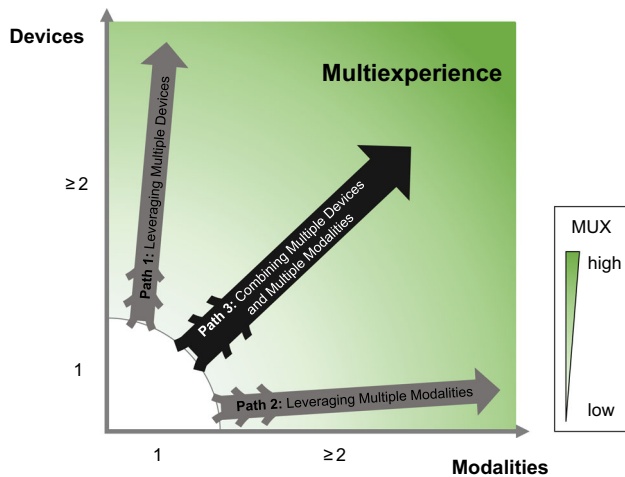


Fig. 1 Conceptual framework of guiding paths toward multiexperience

resulting from the use and/or anticipated use of an IS that leverages multiple devices and/or multiple modalities. As such, it combines the terms *multi* (i.e., multi-device, multimodality) and (user) *experience*¹ to account for the increasing prevalence of more than a single device and/or modality within an IS. New devices and advanced interaction modalities have certainly made scenarios a reality that belonged in the realm of science fiction just a few years ago (e.g., talking to an ERP system or wearing smart glasses that display information directly in the field-of-view). However, since both the IS and HCI field have a long tradition of investigating how users interact with IS across different devices and modalities and how to design IS with multiple devices and/or modalities, we argue that MUX is not an entirely new phenomenon. What is new is that the sheer number of available devices and modalities today provides a greater opportunity and challenge to better meet users' needs and preferences when they interact with an IS. In addition, many devices (e.g., virtual reality headsets) and modalities (e.g., speech interaction) are now beginning to reach a level of maturity that allows widespread application.

To shed a more nuanced light on the concept of MUX, we propose and describe a conceptual framework of guiding paths toward MUX. As depicted in Fig. 1, the framework conceptualizes MUX based on the two axes of devices and modalities, illustrating the shift from single to multiple devices and/or modalities. Drawing on previous research on multi-device and multimodal IS, we propose three paths toward MUX that differ not only in their reliance on multiple devices and/or multiple modalities, but also in their prevalence in prior literature. They are

¹ ISO 9241 defines user experience as the “user’s perceptions and responses that result from the use and/or anticipated use of an interactive system” (ISO 2010).

intended to serve as starting points for individuals and organizations who seek to proceed on the path toward MUX. In this spirit, our framework is meant to assist; not to constrain or suggest that other paths are not possible. There are four important points to be highlighted. First, our framework suggests that MUX is not achievable when there is only a single device and a single modality. Fundamentally, MUX requires at least two devices or two modalities, but not necessarily both (i.e., multiple devices and multiple modalities). Although many of today’s devices technically support multiple modalities, applications running on these devices do not always capitalize on this potential. For example, a smartphone-only application that only supports touch input and visual output would not be able to achieve MUX. Second, our framework allows for variance in MUX. Similar to the use of the UX concept, MUX can vary from low to high. Just because another device is supported or another modality is added does not automatically imply that MUX has improved. For example, most websites today can be accessed from multiple devices, but a website that offers the exact same layout and content across all devices would achieve rather low MUX (e.g., because text could be difficult to read on a smartphone or large images could result in slow loading times). In contrast, higher MUX could be achieved when the website is optimized for each device (i.e., using a responsive design) or dedicated apps are developed for smartphones and tablets. Consequently, the extent to which MUX is achieved also depends on how devices and/or modalities are integrated and allow users to transition from one device or modality to another during use. Third, there are different entry points to each of the three paths. In this sense, our framework is not bound to a specific sequence or set of devices and/or modalities to achieve higher MUX. For example, one company might enter the path to MUX by adding a mobile app to run alongside their website, whereas another might choose a very different path by adding speech interaction capabilities to their ERP system. Finally, our framework suggests that MUX can be achieved by moving along one axis – either vertically along Path 1 (devices) or horizontally along Path 2 (modalities) – or along both axis simultaneously. However, as we explain in the next sections, the greatest potential lies in Path 3 that leverages both multiple devices *and* multiple modalities rather than focusing on one dimension alone.

3.1 Path 1: Leveraging Multiple Devices

Technological development has brought – and continues to bring – a constant stream of new devices to the market. Therefore, taking advantage of these new devices and the opportunities that they offer is a common path toward MUX. For example, starting in the 1990s, mobile devices

became increasingly popular and today they are an essential part of our everyday lives. Many companies have developed mobile apps to complement existing desktop applications. However, particularly at the beginning, this often resulted in merely migrating functionality, content, and design from existing applications to mobile devices without taking their specific characteristics, such as smaller screens and limited keyboard input, into account (Levin 2014). Similar trends can be observed for AR/VR devices with applications that try to faithfully recreate existing functionality from mobile or desktop applications without taking the main advantage of AR/VR – not being bound to characteristics of physical reality – into account (Berke-meier et al. 2019; Wohlgenannt et al. 2020). These observations indicate that leveraging multiple devices can be a viable path toward MUX. However, when such efforts are based on a limited understanding of technology characteristics, task characteristics, and user needs, it will be rather difficult to achieve higher MUX. The following two examples may serve to illustrate past and current efforts on this path toward MUX.

3.1.1 Example #1: Mobile Banking Apps

With the advent of the internet in the 1990s, banks started offering online banking to supplement traditional offline (e.g., ATM, local branch) and phone banking. Customers with a PC connected to the internet could access their accounts and conduct financial transactions through the bank's website. A few years later, when the first cell phones came out, banks launched the first mobile banking services via SMS. However, mobile banking only became an important banking channel after smartphones were introduced at the end of the 2000s. Today, most banks offer native mobile banking apps that enable customers to access their bank accounts through smartphones and tablets in order to conduct a range of financial transactions, including balance checks, fund transfers, and stock trading. However, customers' usage of mobile banking apps often remains either rudimentary (e.g., only checking balances) or lacking altogether (Crowe et al. 2017). Hoehle et al. (2017) provide the example of a bank that spent EUR 300,000 on designing a mobile banking app that was only used by a handful of customers. Research suggests that customers' usage patterns of mobile banking apps are related to technology characteristics (Hoehle et al. 2017; Kim et al. 2009). Some tasks (e.g., more complex financial transactions) may be too difficult to perform on a mobile device due to its small screen and on-screen keyboard. In contrast, the same task may be much easier to perform on a laptop or desktop computer with a physical keyboard and a larger screen. Consequently, when leveraging multiple devices, it is also important to develop a thorough understanding of

technology characteristics (e.g., screen size) and task characteristics (e.g., simple vs. complex) in order to balance the strengths and weaknesses of each device.

3.1.2 Example #2: Augmented and Virtual Reality in E-Commerce

The gaming industry is considered the pioneer in the use of AR and VR (Wohlgenannt et al. 2020). However, AR and VR applications that can, for example, be experienced via head-mounted displays, smart glasses, or smartphones are increasingly employed by e-commerce providers as well (Wedel et al. 2020). Their main goal is to overcome e-commerce's inherent limitation "that online consumers can only passively understand the product information but cannot touch and feel the product" (Tarafdar et al. 2019, p. 1). AR and VR applications can allow consumers to evaluate products in real scale and from different angles (Peukert et al. 2019). For example, to complement existing online shopping experiences via their website and mobile apps, the Swedish furniture company IKEA has developed an AR application that enables consumers to view furniture in real size, from different angles (360° view), and at the intended place (Ozturkcan 2020). Furthermore, IKEA provides different VR applications that help consumers to increase the imagination of product arrangements and encourage co-creation with others. Similarly, Europe's largest retailer for consumer electronics (the MediaMarktSaturn Retail Group) offers a holistic VR shopping environment, Virtual SATURN, encompassing several products from their online shop. Nevertheless, there are also examples of AR and VR applications that are just standalone "gimmicks" – either to serve as marketing tools or as a means to gain experience with this novel technology (Peukert et al. 2019). Consequently, it is important to reflect on the use of AR and VR for tasks that have little additional benefit when compared to physical reality (Steffen et al. 2019) and explore how to integrate AR and VR applications with existing applications offered on other devices in a way that generates substantial added value for consumers.

3.2 Path 2: Leveraging Multiple Modalities

Humans interact with the world through multiple senses and effectors. However, most IS have traditionally focused on unimodal interaction (i.e., a single input and output modality), such as providing visual output on a screen with a keyboard for input. In recent years, system designers have begun to complement the more traditional modalities, such as mouse, keyboard, and touch, with more advanced modalities such as speech or mid-air hand gestures. For example, Apple's Siri allows users to perform various

commands via speech (e.g., setting an alarm or creating a to-do list), which traditionally had to be made via key-strokes or touching buttons. Similar virtual assistants are introduced to the workplace to assist in work-related tasks (Mirbabaie et al. 2021; Seeber et al. 2020). Furthermore, e-commerce providers are increasingly complementing touch- and mouse-based interaction with gesture-based interaction (i.e., reaching, pointing, and manipulating products using hand movements in the air) in order to provide a more natural interaction experience (Liu et al. 2019). However, multimodal interaction capabilities alone do not automatically result in a better or more natural interaction with a system. The number one myth about multimodality is that “if you build a multimodal system, users will interact multimodally” (Oviatt 1999). Whether or not users interact multimodally depends upon many factors, including the nature of the task, the current environment, as well as the user’s individual expectations, experience, and needs. Moreover, different modalities vary in the degree to which they are capable of transmitting similar information (Oviatt 1999). For example, comprehensive results from data analyses in a BI&A system are rather difficult to communicate to users via speech output, while the same information can be easily conveyed using visual output in the form of a graph or chart. Therefore, simply replicating the functionality of one modality in another modality is unlikely to play to the particular strengths of each modality. This is particularly important when multiple modalities can be used simultaneously or sequentially (e.g., pointing at an object and then speaking a command). Taken together, leveraging multiple modalities can be another viable path toward MUX. However, providing multimodal capabilities alone is not sufficient to realize the benefits of having more than one modality available. A thorough understanding of the unique strengths and weaknesses of each modality, the nature of tasks, and user needs is also required to achieve higher MUX. The following two examples may serve to illustrate past and current efforts on this path toward MUX.

3.2.1 Example #1: From Smart Speakers to Smart Displays

Looking at the recent history of smart speakers, it is interesting to observe how they have evolved from smart speakers to smart displays. While smart speakers (e.g., Amazon Echo, Google Home) offer only one modality – speech – for input and output, users can interact multimodally with smart displays (e.g., Amazon Echo Show, Google Nest Hub) because they combine speech interaction with a touchscreen display. For example, users can provide input via speech (e.g., “Alexa, order toilet paper”) and touch (e.g., selecting a product by touching a button on

the screen) and receive both speech output (e.g., “Here are some options for toilet paper”) and visual output (e.g., different products with names, images, and prices). In contrast, when a user provides the same speech input to a smart speaker (i.e., “Alexa, order toilet paper”), it would respond with something like “The top choice for toilet paper is (product name). It costs (product price) euro in total. Would you like me to order it?”. Since it is difficult to present several alternatives via speech output, the smart speaker selects a “top choice”, for example based on the users’ shopping history, and asks them if they want to make the purchase. However, while this product selection process can increase efficiency, it also comes at the cost of transparency and control (Rzepka et al. 2020b). A lack of transparency and control may be less critical for routine tasks, such as playing music or getting weather updates, but play an important role in high involvement tasks (e.g., purchase decisions). Consequently, smart displays try to combine the best of both worlds by leveraging multiple modalities: efficiency via speech in-/output and transparency by augmenting speech output with visual information.

3.2.2 Example #2: Multimodality in Augmented and Virtual Reality

A fundamental characteristic of AR and VR applications is their extensiveness – i.e., “the range of sensory modalities accommodated” (Slater and Wilbur 1997, p. 605). AR/VR applications not only differ in the number of modalities they offer but also in the extent to which these modalities are stimulated. Although head mounted displays (HMDs) with visual output represent the main category of output devices, additional input and output modalities are increasingly integrated into AR/VR devices and, in turn, leveraged by applications (for a comprehensive overview of input and output devices, see Anthes et al. 2016). For example, many HMDs already provide audio output and, in the future, this modality may be complemented with haptic feedback ranging from controller vibrations to realistic force feedback in order to make forces originating from virtual objects perceptible (e.g., HaptX Gloves or Tesla-suit). The latest HMDs are equipped with various sensors that provide the necessary hardware components for multimodal interactions. For example, Facebook’s Oculus Quest 2 is able to capture hand movements via the built-in external cameras and understand speech commands. Therefore, users can control apps through gesture- and speech-based interaction as well. These functionalities can – when supported by an application – make controllers obsolete. For example, YouTube’s VR app can be controlled with only hand gestures. Microsoft’s HoloLens 2 goes one step further and – in addition to gestures and

speech – allows gaze-based interactions via integrated eye-tracking technology. As a result, natural interactions are possible even in hands-free scenarios, such as picking tasks in logistics, or remote assistance use cases at the workplace where gloves must be worn or fingers become dirty.

3.3 Path 3: Combining Multiple Devices and Multiple Modalities

The final path in our conceptual framework results from the conflation of both previously described paths. This path can be regarded as the logical next step in the efforts toward MUX because it seeks to leverage both multiple devices *and* multiple modalities. While researchers and practitioners have traditionally focused their efforts on either devices or modalities, recent years have seen an increased interest in the combination of both elements. For example, Domino's AnyWare platform allows customers to order pizza in 15 different ways using various devices – laptops, smartphones, smart speakers, smart watches, smart TVs, and even cars – and different modalities – mouse and keyboard, touch, and speech (Domino's 2021). However, as the example indicates, this path can quickly become complex and unmanageable because of the sheer number of available devices and modalities. The constant stream of new devices and mature modalities leads to an overwhelming number of possible combinations (i.e., number of devices times the number of modalities supported by each device). Therefore, the key challenge on this path is to choose wisely among the numerous possibilities and identify those combinations that provide the greatest benefit to users. To tackle the increased complexity, it is essential to develop a holistic understanding of the interplay between devices and modalities (e.g., strengths and weaknesses), tasks as well as user needs and preferences. Despite these challenges, the plethora of options also opens up the opportunity to balance the relative strengths and weaknesses of different devices and modalities in order to better meet users' needs overall. Hence, we argue that this path is the one with the greatest potential for both improving user interaction with IS and making significant contributions to research. The following two examples may serve to illustrate past and current efforts on this path toward MUX.

3.3.1 Example #1: Multiexperience in Enterprise Resource Planning (ERP) Systems

Traditionally, ERP systems have focused on a consistent graphical user interface that provides visual output and allows users to interact with the system via mouse and keyboard (Klaus et al. 2000). Most employees access their company's ERP system when they are in their office at a

desk equipped with a PC, mouse, and keyboard. Given the complex information that is conveyed to users through transaction screens, tables, and reports as well the nature of work, the workplace setting will likely continue to be the dominant context of use. However, with the growing popularity of mobile devices in the 2000s, software vendors started to enable mobile access to ERP systems without requiring a local ERP client (Markus et al. 2000). Today, many vendors offer mobile applications for smartphones and tablets (e.g., Sage Mobile Sales app) and provide platforms or frameworks to allow customers to develop their own mobile applications (e.g., Oracle Mobile Application Framework). Recently, ERP systems have also started integrating speech interaction via written or spoken language to complement existing modalities of the traditional graphical user interfaces (vom Brocke et al. 2018). For example, in 2017, SAP launched CoPilot, a digital assistant integrated into the SAP Fiori user interface that can be operated via speech (SAP 2017). Instead of using mouse and keyboard to enter transaction codes, users can also speak or write natural language commands to navigate the interface and perform routine tasks (e.g., "create a sales order for customer X"). Moreover, CoPilot can not only be used from inside an SAP application, but also comes as a standalone mobile application for smartphones and tablets. As a result, users are able to switch seamlessly between desktop and mobile devices and between traditional and speech modalities according to their individual preferences and the characteristics of the task at hand. At their desks, users may prefer to interact with the Fiori user interface via mouse and keyboard, while only occasionally using the CoPilot to enter natural language commands for specific transactions. At home and on the go, they may favor the CoPilot mobile app and interact with it via speech commands, for example, to get a quick overview of current sales and inventory levels for an upcoming meeting. As a result, more efficient and intuitive interactions with an ERP system are possible through combining multiple devices and multiple devices. Therefore, users can not only select the most suitable device and modality for their current task, but also perform a task seamlessly across multiple devices and multiple modalities.

3.3.2 Example #2: Multiexperience in Business Intelligence and Analytics (BI&A) Systems

Important business decisions need to be made by individuals, teams, or groups in many places – at the desk, in meetings, or in the field. To support data-driven decision-making in all of these situations, BI&A systems have evolved from traditional desktop applications for expert users to flexible systems that leverage both multiple devices and multiple modalities to accommodate a wide range

of users (Chen et al. 2012). Today, decision makers can use BI&A systems on their smart phones or tablets while on the go (Power 2013). In meetings, cross-functional teams can make decisions together by collaboratively interacting with BI&A systems on large interactive screens (e.g., Microsoft's Surface Hub) (Ruoff and Gnewuch 2021). To facilitate transparent interaction, particularly for non-expert users, BI&A systems increasingly support multiple modalities. For example, Tableau's Ask Data feature helps users visualize and analyze data by asking a question in natural language. Moreover, mid-air hand gestures have been found to facilitate the collaborative analysis of complex data in BI&A systems (Butscher et al. 2018). Bringing both trends together, BI&A systems increasingly try to combine multiple devices and multiple modalities to better meet decision makers' needs and preferences when performing different tasks. At their desks, individual decision makers may use mouse and keyboard and leverage the large screen of their PCs to perform complex data analyses tasks and prepare detailed management reports. In meetings, teams of decision makers may use a combination of speech- and touch-based interaction to perform ad-hoc analyses on large interactive screens. In the field, particularly when hands-free operation is required, individuals or teams may use speech or gestures to analyze data on the fly. As a result, faster and more intuitive interactions with BI&A systems are possible through combining multiple devices and multiple modalities. Moreover, these BI&A systems put the human at the center because they allow users to choose the device and modality most suited for the characteristics of the task at hand and switch between them in accordance with their individual preferences.

4 Future Research Directions

This catchword sheds light on the concept of MUX and proposes a framework with two dimensions – devices and modalities – and three different guiding paths toward MUX. Drawing on illustrative examples for each path, we embed the concept of MUX into existing streams of research in IS and HCI, explain benefits and challenges of each path, and offer practical guidance for moving along these paths toward MUX. While substantial research has been conducted on the first two paths toward MUX, less attention has been paid to the third path, which seeks to leverage the combination of multiple devices and multiple modalities. Therefore, many promising research questions remain and the BI&A community is well suited to address them. In the following, we suggest four areas for future research on MUX. Table 3 summarizes the identified future research directions and provides illustrative research questions.

First, while this catchword represents a valuable step toward a better understanding of MUX, our conceptualization could benefit from further refinement. Our framework of paths toward MUX provides sufficient structure to guide future research, but also leaves ample room for further exploration of the nature of MUX and the interplay between devices and modalities. For example, future research could systematically identify and classify the many different combinations of devices and modalities that can be used for MUX (e.g., in the form of taxonomies or morphological boxes). Another vital step would be to operationalize the concept of MUX and develop suitable measurement instruments. These instruments would be equally useful for researchers who seek to empirically evaluate MUX and practitioners who want to assess their software products' utility. Existing measurement instruments for usability and UX, such as the system usability scale (Brooke 1996) and the user experience questionnaire (Laugwitz et al. 2008), could serve as a suitable starting point. Additionally, future research could identify ways to measure MUX objectively using behavioral data, such as interaction logs across devices and modalities, to complement self-report measures.

Second, another promising direction for future research is the empirical investigation of MUX. From an individual perspective, such work could, for example, examine whether and how MUX influences the adoption and use of IS. Since MUX involves utilitarian and hedonic aspects, studies should investigate both instrumental (e.g., better performance) and experiential outcomes (e.g., enjoyment) as well as potential trade-offs and synergies between them. Given the important role of contextual factors in MUX, future research is also needed to better understand how users behave in different contexts (e.g., at work, at home, while riding on a subway), when performing different tasks (e.g., information search, online transactions, entertainment), and when switching between different contexts and tasks. Moreover, future studies should consider how individual differences, such as demographics, personality characteristics, and preferences for different devices or modalities, affect MUX over time and whether users can be classified into different MUX user types (e.g., mobile-first users, keyboard- or touch-only users). Finally, from an organizational perspective, empirical investigations could attempt to shed light on whether and when investments into MUX (e.g., developing an AR app) pay off and how organizations can find and achieve an "optimal" level of MUX.

Third, numerous opportunities exist for future research to deliver new design knowledge through building and/or evaluating innovative MUX artifacts. Such research could follow a design-oriented behavioral research approach (Maedche et al. 2021) to observe and analyze existing

Table 3 Suggested directions and questions for a research agenda on multiexperience (MUX)

Research directions	Illustrative research questions
Conceptualization and Operationalization of MUX	<ul style="list-style-type: none"> •How to create a MUX taxonomy or classification based on the different characteristics and combinations of devices and modalities? •Can existing measurement instruments for usability and UX adequately capture MUX or is the development of specific MUX scales necessary? Can we establish a MUX benchmark (in analogy to SUS)? •How can MUX be objectively measured using behavioral data about user interactions with different devices and modalities?
Empirical Investigations of MUX	<ul style="list-style-type: none"> •How does MUX affect users' perceptions and evaluations of an IS before, during, and after use? Does MUX lead to synergies or trade-offs between instrumental and experiential outcomes? •How do different contextual factors (e.g., location), task characteristics (e.g., complexity), and individual differences (e.g., demographics) impact MUX? •Can users be classified into different MUX user types (e.g., mobile-first users, keyboard- or touch-only users)? Do different user types prefer different manifestations of MUX? •How to assess the potential gap between the devices and modalities supported by an IS and their actual use? •How does MUX evolve over time? To what extent does training or experience influence whether users take advantage of MUX capabilities instead of relying on a single device and modality? •How can organizations balance the costs and benefits of increasing their level of MUX? Can there be an optimal number of devices and modalities? Is there something like "MUX maturity"?
Designing Innovative MUX Artifacts	<ul style="list-style-type: none"> •Which fundamental design principles and theories should guide the design of MUX artifacts to effectively combine multiple devices and modalities? •To what extent can existing design knowledge for single-device and unimodal artifacts be reused for the design of MUX artifacts? •How to design MUX artifacts that are capable of adapting to individual users' preferences and behavior over time? <p>How to enable MUX artifacts to automatically change or recommend modalities and/or devices according to the users' current context?</p> <ul style="list-style-type: none"> •How can MUX artifacts be designed to allow users to effortlessly transition from one device or modality to another one after a failure or breakdown?
Methods and Tool Support for MUX Development, Implementation, and Management	<ul style="list-style-type: none"> •How to develop methods and tools to support researchers and practitioners in identifying the most suitable and promising combinations of devices and modalities for a particular purpose? •How to provide methodological guidance on implementing MUX into an existing IT landscape? What technical prerequisites need to be in place? •How to develop platform-independent tools that enable individuals with different backgrounds to design for MUX together? •How to create a domain-specific language to facilitate the integration of and communication across different devices and modalities?

MUX artifacts (e.g., SAP CoPilot, Tableau Ask Data, Amazon's Echo devices) or a design science research (DSR) approach to build new MUX artifacts that tackle important real-world problems. Moreover, a better understanding of whether and how existing design knowledge for single-device and unimodal artifacts can be reused for the design of MUX artifacts would be beneficial. Of particular interest could be to provide design knowledge on how to effectively combine devices and modalities to be able to adapt the MUX to individual users and their changing needs over time. For example, future research could investigate the design of IS that automatically change or recommend modalities and devices according to the users' current needs. Research in this area may also profit from setting up collaborations with researchers from other fields such as computer science.

Finally, future research should aim to provide methods and tools that support the development, implementation, and management of MUX. While existing methods and tools from areas, such as human-centered design, UX, and software engineering, may be used as a starting point, it is evident that handling the complexity of MUX – resulting from the large number of possible combinations of devices and modalities – requires a new set of methods and tools. For example, future research could develop methodological guidance to help researchers and practitioners choose among the plethora of options in order to identify the most suitable and promising combinations of devices and modalities for a particular purpose. Similarly, methodological guidance on implementing MUX in an existing IT landscape would be valuable. Finally, while many software vendors offer their own MUX development platforms

(Gartner 2021), future research could provide platform-independent tools that empower everyone to design for MUX, regardless of whether they use commercial software packages or their own software stack.

Overall, we believe that this catchword offers a fresh perspective on MUX and opens up manifold opportunities for future research. MUX has not only been an important theme in prior IS research, but current trends and the constant technological advancement also indicate that it will continue to be so for the foreseeable future. At the same time, it is clear that the multitude of ways in which devices and modalities can be combined adds another layer of complexity to understanding the interplay between user, task, and technology. For example, it is difficult enough to design a unimodal artifact for one specific device or to rigorously examine how users interact with an IS on a single device using one modality. Going forward, there is little doubt that these difficulties will increase as the number of available devices and mature modalities continues to grow. Given the background, interests, and skills of BISE researchers, we are convinced that the BISE community is well positioned to both address the challenges and take advantage of the opportunities for future research on MUX, and we invite fellow researchers to contribute to this exciting research stream.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anthes C, Garcia-Hernandez RJ, Wiedemann M, Kranzlmüller D (2016) State of the art of virtual reality technology. In: Proceedings of the 2016 IEEE aerospace conference. <https://doi.org/10.1109/AERO.2016.7500674>
- Berkemeier L, Zobel B, Werning S, Ickerott I, Thomas O (2019) Engineering of augmented reality-based information systems. *Bus Inf Syst Eng* 61(1):67–89. <https://doi.org/10.1007/s12599-019-00575-6>
- Bolt RA (1980) “Put-that-there”: voice and gesture at the graphics interface. *ACM SIGGRAPH Comput Graph* 14(3):262–270. <https://doi.org/10.1145/965105.807503>
- Brooke J (1996) SUS: a “quick and dirty” usability scale. In: Jordan P et al (eds) Usability evaluation in industry. Taylor & Francis, Milton Park, pp 189–194
- Brudy F, Holz C, Rädle R, Wu C-J, Houben S, Klokmoose CN, Marquardt N (2019) Cross-device taxonomy: survey, opportunities and challenges of interactions spanning across multiple devices. In: Proceedings of the 2019 CHI conference on human factors in computing systems. <https://doi.org/10.1145/3290605.3300792>
- Butscher S, Hubenschmid S, Müller J, Fuchs J, Reiterer H (2018) Clusters, trends, and outliers. In: Proceedings of the 2018 CHI conference on human factors in computing systems. <https://doi.org/10.1145/3173574.3173664>
- Chen H, Chiang RHL, Storey VC (2012) Business intelligence and analytics: from big data to big impact. *MIS Q* 36(4):1165–1188. <https://doi.org/10.2307/41703503>
- Crowe M, Tavilla E, McGuire B (2017) Mobile banking and payment practices of U.S. financial institutions: 2016 mobile financial services survey results from FIs in seven Federal Reserve districts. Federal Reserve Bank of Boston. <https://www.bostonfed.org/publications/mobile-banking-and-payment-surveys/mobile-banking-and-payment-practices-of-us-financial-institutions.aspx>. Accessed 18 Feb 2021
- Dearman D, Pierce JS (2008) “It’s on my other computer!”: computing with multiple devices. In: Proceeding of the 2008 CHI conference on human factors in computing systems, pp 767–776. <https://doi.org/10.1145/1357054.1357177>
- Diederich S, Brendel AB, Kolbe LM (2020) Designing anthropomorphic enterprise conversational agents. *Bus Inf Syst Eng* 62(3):193–209. <https://doi.org/10.1007/s12599-020-00639-y>
- Domino’s (2021) Domino’s anywhere. <https://anywhere.dominos.com/>. Accessed 4 Mar 2021
- Dong T, Churchill EF, Nichols J (2016) Understanding the challenges of designing and developing multi-device experiences. In: Proceedings of the 2016 ACM conference on designing interactive systems, pp 62–72. <https://doi.org/10.1145/2901790.2901851>
- Gartner (2019) Gartner top 10 strategic technology trends for 2020. <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/>. Accessed 17 Sep 2020
- Gartner (2021) Multiexperience development platforms (mxdp) reviews 2021. <https://www.gartner.com/reviews/market/multiexperience-development-platforms>. Accessed 12 Mar 2021
- GlobalWebIndex (2020) GlobalWebIndex’s flagship report on device ownership and usage. <https://www.globalwebindex.com/reports/device>. Accessed 3 Mar 2021
- Hoehle H, Kude T, Huff S, Popp K (2017) Service-channel fit conceptualization and instrument development. *Bus Inf Syst Eng* 59(2):97–110. <https://doi.org/10.1007/s12599-015-0415-z>
- ISO (2010) ISO 9241-210:2010: Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. International Organization for Standardization
- Jaimes A, Sebe N (2007) Multimodal human-computer interaction: a survey. *Comput vis Image Underst* 108(1–2):116–134. <https://doi.org/10.1016/j.cviu.2006.10.019>
- Kim G, Shin B, Lee HG (2009) Understanding dynamics between initial trust and usage intentions of mobile banking. *Inf Syst J* 19(3):283–311. <https://doi.org/10.1111/j.1365-2575.2007.00269.x>
- Klaus H, Rosemann M, Gable G (2000) What is ERP? *Inf Syst Front* 2(2):141–162
- Köffer S, Ortbach K, Junglas I, Niehaves B, Harris J (2015) Innovation through BYOD? The influence of IT consumerization on individual IT innovation behavior. *Bus Inf Syst Eng* 57(6):363–375. <https://doi.org/10.1007/s12599-015-0387-z>
- Laugwitz B, Held T, Schrepp M (2008) Construction and evaluation of a user experience questionnaire. In: Holzinger A (ed) HCI and usability for education and work. Springer, Heidelberg, pp 63–76. https://doi.org/10.1007/978-3-540-89350-9_6
- Lee H-K, Suh K-S, Benbasat I (2001) Effects of task-modality fit on user performance. *Decis Support Syst* 32(1):27–40. [https://doi.org/10.1016/S0167-9236\(01\)00098-7](https://doi.org/10.1016/S0167-9236(01)00098-7)

- Levin M (2014) Designing multi-device experiences: an ecosystem approach to user experiences across devices. O'Reilly Media, Sebastopol
- Li NL, Zhang P (2005) The intellectual development of human-computer interaction research: a critical assessment of the MIS literature (1990–2002). *J Assoc Inf Syst* 6(11):227–292. <https://doi.org/10.17705/1jais.00070>
- Liu Y, Jiang Z, Chan HC (2019) Touching products virtually: facilitating consumer mental imagery with gesture control and visual presentation. *J Manag Inf Syst* 36(3):823–854. <https://doi.org/10.1080/07421222.2019.1628901>
- Lyytinen K, Yoo Y (2002) Research commentary: the next wave of nomadic computing. *Inf Syst Res* 13(4):377–388. <https://doi.org/10.1287/isre.13.4.377.75>
- Maedche A, Gregor S, Parsons J (2021) Mapping design contributions in information systems research: the design research activity framework. *Commun Assoc Inf Syst* 49(1):355–378. <https://doi.org/10.17705/1CAIS.04914>
- Markus ML, Petrie D, Axline S (2000) Bucking the trends: what the future may hold for ERP packages. *Inf Syst Front* 2(2):181–193. <https://doi.org/10.1023/A:1026548107263>
- Mirbabaie M, Stieglitz S, Brünker F, Hofeditz L, Ross B, Jun FNR (2021) Understanding collaboration with virtual assistants – the role of social identity and the extended self. *Bus Inf Syst Eng* 63(1):21–37. <https://doi.org/10.1007/s12599-020-00672-x>
- Netflix (2013) Netflix quick guide: how to continue watching on a different device. <https://www.youtube.com/watch?v=67WFBbBpeSs>. Accessed 21 Dec 2020
- Nigay L, Coutaz J (1993) A design space for multimodal systems: concurrent processing and data fusion. In: Proceedings of the 1993 CHI conference on human factors in computing systems, pp 172–178. <https://doi.org/10.1145/169059.169143>
- Obrist M, Velasco C, Vi C, Ranasinghe N, Israr A, Cheok A, Spence C, Gopalakrishnakone P (2016) Sensing the Future of HCI. *Interact* 23(5):40–49. <https://doi.org/10.1145/2973568>
- Oviatt S (1999) Ten myths of multimodal interaction. *Commun ACM* 42(11):74–81. <https://doi.org/10.1145/319382.319398>
- Ozturkcan S (2020) Service innovation: using augmented reality in the IKEA Place app. *J Inf Technol Teach Cases*. <https://doi.org/10.1177/2043886920947110>
- Peukert C, Pfeiffer J, Meißner M, Pfeiffer T, Weinhardt C (2019) Shopping in virtual reality stores: the influence of immersion on system adoption. *J Manag Inf Syst* 36(3):755–788. <https://doi.org/10.1080/07421222.2019.1628889>
- Power DJ (2013) Mobile decision support and business intelligence: an overview. *J Decis Syst* 22(1):4–9. <https://doi.org/10.1080/12460125.2012.760267>
- Reeves LM, Lai J, Larson JA, Oviatt S, Balaji TS, Buisine S, Collings P, Cohen P, Kraal B, Martin J-C, McTear M, Raman T, Stanney KM, Su H, Wang QY (2004) Guidelines for multimodal user interface design. *Commun ACM* 47(1):57–59. <https://doi.org/10.1145/962081.962106>
- Rekimoto J, Saitoh M (1999) Augmented surfaces: a spatially continuous work space for hybrid computing environments. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '99), pp 378–385. <https://doi.org/10.1145/302979.303113>
- Rekimoto J (1997) Pick-and-drop: a direct manipulation technique for multiple computer environments. In: Proceedings of the 10th annual ACM symposium on user interface software and technology (UIST '97), pp 31–39. <https://doi.org/10.1145/263407.263505>
- Ruoff M, Gnewuch U (2021) Designing multimodal BI&A systems for co-located team interactions. In: Proceedings of the 29th European conference on information systems (ECIS 2021), a virtual AIS conference.
- Rzepka C, Berger B, Hess T (2020a) Is it a match? Examining the fit between conversational interaction modalities and task characteristics. In: Proceedings of the 41st international conference on information systems, India
- Rzepka C, Berger B, Hess T (2020b) Why another customer channel? Consumers' perceived benefits and costs of voice commerce. In: Proceedings of the 53rd Hawaii international conference on system sciences, pp 4079–4088
- SAP (2017) Hey ... I am talking to you! <https://experience.sap.com/news/hey-i-am-talking-to-you/>. Accessed 2 Mar 2021
- Seeber I, Bittner E, Briggs RO, de Vreede T, de Vreede G-J, Elkins A, Maier R, Merz AB, Oeste-Reiß S, Randrup N, Schwabe G, Söllner M (2020) Machines as teammates: a research agenda on AI in team collaboration. *Inf Manag* 57(2):103174. <https://doi.org/10.1016/j.im.2019.103174>
- Slater M, Wilbur S (1997) A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence Teleoper Virt Environ* 6(6):603–616. <https://doi.org/10.1162/pres.1997.6.6.603>
- Steffen JH, Gaskin JE, Meservy TO, Jenkins JL, Wolman I (2019) Framework of affordances for virtual reality and augmented reality. *J Manag Inf Syst* 36(3):683–729. <https://doi.org/10.1080/07421222.2019.1628877>
- Suh K-S, Lee YE (2005) The effects of virtual reality on consumer learning: an empirical investigation. *MIS Q* 29(4):673–697. <https://doi.org/10.2307/25148705>
- Tarafdar P, Leung A, Leung A (2019) Impact of immersive interface design on consumer perceptions during online product presentation. In: Proceedings of the 40th international conference on information systems, Munich
- Turk M (2014) Multimodal interaction: a review. *Pattern Recogn Lett* 36(1):189–195. <https://doi.org/10.1016/j.patrec.2013.07.003>
- vom Brocke J, Maaß W, Buxmann P, Maedche A, Leimeister JM, Pecht G (2018) Future work and enterprise systems. *Bus Inf Syst Eng* 60(4):357–366. <https://doi.org/10.1007/s12599-018-0544-2>
- Wedel M, Bigné E, Zhang J (2020) Virtual and augmented reality: advancing research in consumer marketing. *Int J Res Mark* 37(3):443–465. <https://doi.org/10.1016/j.ijresmar.2020.04.004>
- Weiser M (1991) The computer for the 21st century. *Sci Am* 265(3):94–105
- Westcott K, Loucks J, Littmann D, Wilson P, Srivastava S, Ciampa D (2020) Connectivity and mobile trends survey - build it and they will embrace it. The Deloitte Center for Technology, Media & Telecommunications. <https://www2.deloitte.com/us/en/insights/industry/telecommunications/connectivity-mobile-trends-survey.html>. Accessed 22 Dec 2020
- Wohlgenannt I, Simons A, Stieglitz S (2020) Virtual reality. *Bus Inf Syst Eng* 62(5):455–461. <https://doi.org/10.1007/s12599-020-00658-9>
- Zhang P, Li NL, Scialdone MJ, Carey J (2009) The intellectual advancement of human-computer interaction research: a critical assessment of the MIS literature (1990–2008). *AIS Trans Hum Comput Interact* 1(3):55–107