

Received April 7, 2020, accepted April 30, 2020, date of publication May 11, 2020, date of current version May 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2993646

The Why, What, and How of Immersive Experience

CHENYAN ZHANG 

Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway

e-mail: chenyan.zhang@ntnu.no

ABSTRACT Assessing and modelling the quality of immersive experience has become a trending topic in QoE. To achieve this purpose, we would firstly want to know why an immersive experience is important to us, what exactly is defined as an immersive experience, and how we could actually measure an immersive experience. This paper aims to answer these questions. We start from the rationales of designing immersive experience, and then briefly go through the definitions of immersion from the system and user perspectives. We also provide our own definition of immersion that fits into the QoE assessment and measurement paradigm. We continue to discuss the pros and cons of four popular measurements of immersion, namely, the psychometric questionnaires, the continuous subjective measures, the primary or secondary task performance, and the neuro-psycho-physiological methods. We also provide a global view of comparing and evaluating these measurements by profiling them along five quality dimensions. Finally, we postulate, briefly, four novel methods of measuring immersion, linking the definitions and theories of immersion with the measurements we have discussed in this paper.

INDEX TERMS Immersion, extended reality, subjective tests, objective tests, psychometric questionnaire, continuous subjective measures, primary or secondary task performance, neuro-psycho-physiological methods.

I. INTRODUCTION

Despite decades of presence research, the question of precisely framing the phenomenon of immersion or presence still remains unclear. Nevertheless, laboratory experiment paradigms in measuring immersive experience have yet to be well-established and standardized, giving rise to confusions in disambiguating this predominant characteristic in human-computer interaction. Call it engagement, absorption, involvement, transportation, or whatever that represents a mental fixation in a remote venue, immersion in a digital media environment is the sensory and perceptual experience of being physically located in a non-physical, mediated, or simulated virtual environment. In this process, sensations and perceptions related to stimuli and cues outside this virtual environment are totally blocked and cut off to allow for a complete attentional engagement with the virtual world. In other words, one's sensory and perceptual experiences of the real world are completely overridden by their virtual experience, accompanied by a diminished awareness and suspended belief of the physical self and the physical

context. This virtual experience is so overwhelming and complete to the extent that the participant even wants to exert some degree of autonomous control and interactions with the virtual artefacts. One's communicating modalities are all bonded to this virtual environment and integrated together to simulate mental imageries and representations as if this other reality experience would infinitely approximate their real-world experience [1]. With the burgeoning development of Virtual Reality, Augmented Reality and Mixed Reality (VR/AR/MR) technologies, immersion has become a phenomenal experience of user interacting with the virtual environment. And measuring immersion is growing into an increasingly important task since evaluating the quality and user experience of the virtual, augmented or mixed realities depends heavily upon our understanding of immersion, both through the theories that delineate the genus and the practical measurements that encapsulate the heuristics. Assessing and modelling the quality of immersive experience has become a trending topic in QoE, and "designing for immersion" (DfI) is a primary goal that the media producers are fervently chasing. However, investigating the methods of measuring immersion is still a less-trodden path that, although rapid attention has been riveted there, still is perplexing yet

The associate editor coordinating the review of this manuscript and approving it for publication was Tai-Hoon Kim.

perpetuating the development of immersive technologies. Among the many methods of measuring immersion, the two most popular ones are the psychometric questionnaires [2], [3] as a subjective test method and the brain-computer interface (or more broadly scoped, the neuro-psycho-physiological measurements) [4], [5] as an objective test method. These methods either measure the subjective perception and judgment of the immersive environment by seeing immersion as a multi-dimensional construct, or measure the objective biological signals or emotional responses as indicators that are externalized by the user's complex internal mechanisms during the process of immersion. Besides these two popular methods with rather well-defined experiment paradigms, there are other less common methods, for instance, the continuous subjective measures [6], [7] and the primary or secondary task performance [8], [9], that measure immersive experience and they should not remain lost or unnoticed in the vast ocean of immersion literature. These less common methods, though less prominent in our general understanding of immersion measurement, each having its own merits and worthiness that should not be discredited or undermined due to their less visibility. This paper aims to fill this gap, where we not only are introducing the instruments, experimental methods and the underlying theories or rationales of these common and less common measurements, but are also providing our own critiques of them by discussing their advantages and shortcomings, including their suitability or application areas, constraints, etc.

Before we delve deep into the enjoyment of an immersive experience, we would firstly want to know why an immersive experience is important, what exactly is defined as an immersive experience, and how we can actually measure an immersive experience. With these questions in mind, we embark on a hopefully enjoyable and interesting journey in searching for the answers in this paper. We start from the rationales of designing immersive experience, journeying to the definitions of immersion from the system and user perspectives – immersion means either the system engages us or we are captivated by the system. We also provide our own definition of immersion that fits into the QoE assessment and measurement paradigm. And we travel our way through to discussing the pros and cons of four popular methods of measuring immersive experience aforementioned. We also provide a global view of comparing and evaluating these measurements by profiling them along five quality dimensions. Finally, we propose, in a proof-of-concept manner, four blue sky ideas of measuring immersion, linking the definitions and theories of immersion with the measurements we have discussed in this paper. Now, take a deep breath, fasten your seatbelt, and the journey begins! Phew...

II. WHY “IMMERSIVE EXPERIENCE”?

Before the outbreak of digital shockwaves, when speaking of immersive experience, one may easily relate it with film spectatorship. How the emotional resonance between both sides of the screen can be so strong and resounding has always

been, and will still be, a central theme in film and media studies; Or perhaps looking even further afield, imagine the moments when we are deeply touched by a story in a print media, such as reading a novel or newspaper, or get lost before a painting in a museum, or feel nostalgically aroused among a concert; Or dated back furthest, consider the occasions when we were gathering together around a campfire and listening attentively to a folklore story told by our elder generations. Yes, immersive experience emerges and grows alongside the development of storytelling. With the advent of digital storytelling, particularly VR storytelling, i.e. storytelling with virtual, augmented and mixed realities, immersion still remains an indispensable characteristic of user interacting with the virtual environment. Why an immersive experience is important when it comes to designing and experimenting with the virtual environment? We consider the following three aspects as something of the most pertinent:

A. TO FULFILL THE PURPOSE OF XR ENTERTAINMENT

“Extended Reality” (XR) entertainment is a concept incorporating virtual reality, augmented reality and mixed reality, and suggests a simulated and imaginary reality that is dictated and extended by our physical reality, and both the real and the virtual are responsively interacting with each other. Intuitively enough, XR entertainment, by means of its virtuality, has deeply embedded in its innate nature of “immersiveness”. This immersiveness is triggered by the “immersive cues” fabricated by immersive technologies in the physical environment, which, by catering to our perceptual and sensory organs, form the “immersive stimuli” that stimulate the formation of the perception and mental imagery schemata of “immersion”, the illusory embodiment as if the user is physically located in a simulated environment. By taking advantage of the “immersive cues and stimuli”, the XR entertainment allows us to experience an alternate reality that only exists in our imagination, and this gives us the thrills to explore some of the previously uncharted territories, particularly in situations where we let our wildest imagination unwind to face and overcome the phobia and perils which in reality we dare not to encounter or were thwarted from. In addition, immersive experience allows the users to feel, sense, or even interact with or manipulate the new products in a virtual prototype, before they were produced or manufactured to the market. And this lends vast potentials to iterate the product life cycle and appropriate user experience in a cost-effective manner.

B. TO CREATE CLOSE TO REALITY FULL SENSORY SIMULATION

We as human beings are inhabited in a built, urban, or landscaped environment. Considering the complexities of our environments, how and to what extent we are emotionally attached by, feel comfortable or at home with, or have an affinity with the environment we are living in is not easy to quantify or measure. Yet this is a paramount indicator of our quality of life, for instance how we would fully immerse

ourselves in a larger-than-life experiential marketing event. Fortunately, with digital fabrication, we can certainly scale down the built environment to a desktop virtual environment, and imagine and project how we as cyborgs are navigating and interacting within a virtual environment, by which the immersiveness of a real-life event or environment can be simulated. Recent efforts in this area include the “Digital Twins” concept [10], where the everyday functioning of a real city can be simulated and tested against an immersive virtual platform.

C. TO INVESTIGATE THE MIND-BODY PHENOMENON OR PROBLEM

Mind-body phenomenon is, by far, the essential and crucial problem in the understanding of human consciousness and cognition, and thus is the corner-stone that perplexes and perpetuates the developments in many human factor related domains such as Human-Computer Interaction (HCI), Artificial Intelligence (AI), sensory and perceptual experiences (i.e. sensation and perception), creativity and creative cognition, visual aesthetics, phenomenology, and, certainly, Quality of Experience (QoE), where the quality of multimedia service is measured against human experiences and perceptions.

Although dualism contends plausibly that the mind or human consciousness is, at least partly, independent of the body, yet it is still difficult to separate the mind from the body in a living and active organism, and to measure and experiment with either of them under a separated state. Scarcities in methods to achieving this aim have greatly limited our understanding of consciousness, although this certainly would entail an ethical debate. The immersive XR technologies, given their nature to create or induce the transcendental experience of absorbing the mind to actively interact in a virtual space, potentially hold the promise and provide the facility to separate the mind from the body in an ongoing and continuous process. This “out-of-body” experience in the process of immersion has been claimed by many users of XR technologies, and is validated and clearly described in the psychometric immersive instruments [3], for instance “During the story, my body was in the room, but my mind was inside the world created by the story”, or, “When the video ended, I felt like I came back to the ‘real world’ after a journey”, or, “The story came to me and created a new world for me, and the world suddenly disappeared when the video ended”. More precisely, this “out-of-body” experience is even clearly defined and embedded in the very nature of immersion – the sense of “being there”. Loomis [11] contemplates that immersion is a basic state of consciousness, affording the consciousness the capacity to loom beyond the proximity of the body and dwell in a remote site while still remain the cognizant ownership of the body, for a limited time span. In this vein, Sas & O’Hare [12] define immersion as “an imperceptible shifting of focus of consciousness to the proximal stimulus located in a technological mediated or imaginary world.” XR technologies may also, on the other

hand, provide counter-examples to prove that the mind is by no means separable from the body (although this is apparently vis-à-vis the fact of immersion), and in that case dualism might be fallacious. Discoveries in either direction could provide deep insights on demystifying the long-standing mind-body problem, and thus help us gain a better understanding of how the human consciousness and cognition work.

III. SO, WHAT IS AN IMMERSIVE EXPERIENCE?

The varied arrays of definition of “immersion” reflect the fact that it is a multi-facet phenomenon. To make it happen requires the careful design and manufacture from the system side, and the mental readiness and facility from the user side. There are many concepts in describing this sensational experience of consciousness displacement, in particular, immersion and presence are two concepts that delineate the genus. Although previous studies have made efforts to distinguish between immersion and presence, we argue that both immersion and presence can be described as either a proactive function that reflects the design purpose of the multimedia system - “the degree to which a virtual environment submerges the perceptual system of the user [13]”, or a reactive psychological feedback that is perceived and experienced by the user when interacting with the system - “the degree to which users of a virtual environment feel involved with, absorbed in, and engrossed by stimuli from the virtual environment [14].”

From the system side, Turner *et al.* [15] consider immersion as “being positively associated with the degree of technologically-mediated sensory richness that facilitates isolation or decoupling from the real world.” Slater and Wilbur [16] define immersion as being the extent to which a computerized system is capable of offering to the user the illusion of reality at once being: (1) inclusive (with attentional resources and sensory modalities fully engaged with the virtual world, and information from the physical reality completely cut off, suspended or isolated); (2) vast or extensive (with the virtual world providing simulation that accommodates and adapts to a whole repertoire of sensory modalities); (3) surrounding (meaning that a system offers a virtual environment that is panoramic rather than being limited to a narrow field of vision, i.e., a field of view from all virtual directions); and (4) vivid (meaning that the fidelity of the stimuli is sharp and rich to provide high-quality information, content and interface); and (5) matching (with each sensory modality of the stimuli consistent with one another and altogether providing an experience that is congruent to real-life scenarios).

To create this illusion, the system needs to be able to elicit adequate sense of “perceived realness”. An important element of this “perceived realness” is for the technological capability of the system to provide “the perceptual illusion of non-mediation”. Lombard [17] offers an explanation of presence as being in a psychological state or having a subjective perception in which, even if the experience is generated by technology, a part or a totality of the individual’s perception fails to recognize the role of technology at the

time of the virtual experience. By “non-mediation”, presence is defined as requiring the use of technology and resulting in a psychological state in which media users voluntarily suspend the experience of mediation in order to feel a sense of connection with the mediated content they are using (i.e., connection to characters, involvement in the storyline) [18]. Sanchez-Vives and Slater [19] also pointed out that inside the virtual experience, you are at the same time conscious of the “place” and the “events” and simultaneously conscious of that there are no such place or events; however, you still behave and think as if the place were real and the events were happening. As your consciousness of the differences between the real and virtual place and events blurs, the barrier between your mind and the VE diminishes, improving your interaction with the computer-generated world.

From the human user side, immersion evokes a series of parallelly intertwining psychological changes at the attentional, emotional, cognitive, sensory, perceptual and memory level, that synergically mark the uniqueness of this experience. Singer and Witmer [20] describe presence as a perceptual flux that demands the direct attention of the individual. They suggest that presence be based on the interaction between sensory stimulations, environmental stimulations and the internal tendencies of the person. The individual’s psychological perception of presence within a virtual environment is perceived principally as a by-product of the properties of immersion, and as being implicated in the virtual environment. Thus, presence in a virtual environment depends on the degree of attention of the user as they displace themselves in the physical environment. In summary, immersion is a complex, multidimensional perception, formed through an interplay of raw (multi-)sensory data and various cognitive processes – an experience in which attentional factors play crucial roles [21]. Presence in a mediated environment will be enhanced when the environment is immersive and perceptually salient, as well as when attentional selection processes are directed towards the mediated environment, thus allowing the formation of a consistent environmental representation [22]. Presence occurs when more attentional resources are allocated to the computer-mediated environment: “The more attentional resources that a user devotes to stimuli presented by the displays, the greater the identification with the computer-mediated environment and the stronger the sense of presence” [23]. Witmer and Singer [24] further define immersion as a psychological state characterized by the perception of being or feeling “enveloped by”, “included in” or “in interaction with” an environment offering a continuity of various stimulatory experiences. Slater and Steed’s [22], [25] notion of presence is a perceptual mechanism for organizing the incoming stream of sensory data into a coherent environmental gestalt, thus essentially selecting between alternative hypotheses of self-location: ‘I am in this place’ versus ‘I am in that place’.

Other lines of research suggest that immersion is a mental simulation synthesized from the immersive cues fabricated by the system, and this lays the connection between system

factors and the human user factors. Jones [26] defines that immersion is a “response to a mental model of an environment that takes shape in the mind of the individual based upon a combination of cues that originate both externally and internally”. People generate mental representations through physical simulations, situated action, and bodily states. Grounded cognition and learning can occur at various levels of mental processing, taking into account abstract internal representations [27].

Yet other lines of research take immersion as “perspective-taking”, either spatially and/or emotionally [3]. Spatial perspective taking is roughly another word for “embodiment”, suggesting one’s sensorimotor ability to adopt the spatial locus of the virtual character(s), and update their ego-centric frame of reference from the virtual character’s point of view. Emotional perspective taking means empathy, suggesting one’s experience of cognitively identifying and emotionally empathizing with the virtual character. Embodiment (i.e. spatial perspective taking) and empathy (i.e. emotional perspective taking) are two fundamental forms of immersion in the virtual environment. Perspective taking supports the assumption that with the attentional resources drawing to the virtual space by immersive stimuli/cues, several different physio-psychological processes and pathways are activated all at once, such as the sensori /perceptuo-motor responses and/or the cognitive/affective processing, indicating the interactions and synergies of these psychological components in contributing to the overall experience of immersion.

Despite the diversity of these definitions, a consensus points to the fact that immersion is “the pleasurable experience of being transported to an elaborately simulated place” and “the sensation of being surrounded by a completely other reality that takes over all of our attention and our whole perceptual apparatus” [28]. To sum up, if we cannot precisely define a construct, we cannot measure it, and we cannot further understand, improve and develop it. The fundamental understanding of immersion is that it is an optimal mental state when users are interacting with the virtual system, in which several physio-psychological processes and mechanisms contribute in culminating into the transcendental experience of being physically shifted into the virtual space when it is actually only the mind, directed by attention, that is focused on the fabricated virtual space. Here, the mental shift leads to the illusion of a physical one. From this understanding we can conclude that immersion entails the very notion of “embodiment”. Thus our brief definition of immersion in the virtual reality that fits into the QoE assessment paradigm is as follows:

Immersion in a virtual environment is a technology-mediated illusion that, through mimetic system offering priming stimuli and cues, engulfs one’s senses and leads to the alignment of one’s attentional focus to a synthetic yet perceptually authentic reality, by taking the visuo-spatial and emotional perspectives of the virtual agent(s), depending on one’s imaginative facilities and mental dispositions and tendencies.

IV. MEASURING THE QUALITY OF IMMERSIVE EXPERIENCE

It is necessary to distinguish between merely measuring an immersive experience and measuring the quality of immersive experience. And we should also try to distinguish among, using the neuro-psycho-physiological methods to measure and assess: 1) the QoE of a multimedia system; 2) the QoE of an XR application; and 3) the quality of immersive experience of an XR system. In this paper, we are interested in whether immersion can be reliably measured, aligned with how it is defined, in a way that is relevant to QoE. Since immersion is such a multi-dimensional construct, one may argue that it is too broadly-scoped to be operationalized in details. Within the QoE paradigm, measuring the quality of immersive experience allows us to focus on the “perceived” immersiveness of an XR system, by taking account of the experiential factors for instance evaluating the user’s sensory-perceptual experience using psychometric tests or measuring user’s attentional allocation (cognitive load) through primary- or secondary-task performance or neurophysiological means. With that definition of immersion above, we aim to understand, along the “perception” route, that:

- 1) How “perceptually authentic” an XR system can be?
- 2) How and to what extent are our senses engaged or engulfed, in other words, how do we perceive ourselves as being “enveloped by”, “enclosed in” or “in interaction with” the virtual environment?
- 3) Without oscillating between two realities, how are we voluntarily or passively submitting and continuously sustaining our attention to the virtual system, even having in mind that it is technologically mediated or a simulated illusory perception?
- 4) Since immersion is about perspective taking, to what degree the user is spatially immersed and/or emotional immersed, and which one is more immersive? What are the interactions of spatial perspective taking and emotional perspective taking in the process of immersion?

There are many ways of measuring immersion, for instance, subjective measures and objective corroborative measures. When using subjective measures, a participant is asked for a conscious judgment of his/her psychological state/response in relation to the mediated environment. The objective approach to immersion measurement attempts to measure user responses that are produced automatically and without conscious deliberation, but are still sensibly correlated with measurable properties of the medium and/or the content [2].

A. SUBJECTIVE TESTS

1) PSYCHOMETRIC TESTS

Subjective tests are more closely oriented towards measuring the “perceived” immersiveness of the system, and thus are more frequently used as a quality assessment method of immersion. Subjective tests are usually used as pilot studies for more accurate measurements, to obtain ground

truth, identify patterns, and find correlations among various potent and latent components of immersion. Among the many forms of subjective methods, for instance those popular in social science, e.g. retrospective self-report, think aloud protocol, diary research, focus group, ethnographic observation and semi-structured interviews, psychometric questionnaire stands out as the most widely-adopted and effective way of measuring the subjective experience of immersion. In the paradigm of QoE assessment of immersive experience, psychometric scales/ questionnaires can be used as:

- i) comparing and assessing the level of immersiveness across varied contents, devices, and applications;
- ii) investigating the individual differences in the perception of immersion while using a particular application or device;
- iii) testing the usability, user experience and QoE of a system by probing into the immersive experiences of its users;
- iv) identifying and screening for significant traits of immersion using a set of pre-defined parameters and criteria;
- v) sifting instances and converging ideas by mixing and matching the components of immersion as tested to design new and diversified immersive experience.

The psychometric questionnaire method has some advantages in measuring immersive experience: i) It is easy to administer and interpret. The administration of the experiment does not require complicated system setup. And through cluster or factorial analysis, components of immersion are extracted and grouped, and their correlations identified; ii) It is not intrusive to the immersive experience during the experiment process, and thus the results obtained are more truthfully reflect the nature of immersion. iii) It probes into many aspects of immersive experience, thus is more suitable to measure this multi-dimensional phenomenon. And considering that immersion is a multidimensional construct with many neuro-psycho-physiological processes exerting influences to the overall immersive experience at different levels and with different mechanisms and manifestations, and even considering the fact that these factors themselves are inextricably correlated, psychometric tests can study the individual characteristics of these factors without undermining the inextricability of them, while keeping intact of both their individual features and inextricability, all at once in a single study. It would be fairly difficult to achieve this with other methods. iv) It is very suitable to be used for hypothesis testing since every statement in the questionnaire can be a hypothesis, thus facilitating rigorous theoretical development. v) It is a very robust measurement as it points directly to the subjective experience of immersion without reference to a specific content, device, or application, thus it is very suitable for cross-platform comparison and to identify individual differences in the perception of immersion. vi) Because the psychometric tests are administered afterwards by recalling the participant’s overall immersive experience during the experiment, they are not subject to the abrupt idiosyncrasies of the

participant's personal emotional or physiological conditions that are irrelevant to the test stimuli, in this respect their immersive experiences are rather truthfully reported.

The disadvantages of using psychometric questionnaires to measure immersive experience could be:

- i) Due to the fact that psychometric test relies on the retrospective memory recall immediately after watching the video, it would be difficult to track the variations of immersive experiences throughout the whole process of immersion. For the same reason, it is also subject to recency effect and duration neglect. Thus, it might be easy to identify whether a peak immersive experience does happen or not, but when exactly it happens can still be beyond the means of a psychometric test to track and capture.
- ii) Since the items (questions) in the psychometric tests are pre-defined, based on assumptions and results obtained from other qualitative studies, it is difficult to identify patterns and find peculiarities or serendipities outside the scope of these question items, thus limiting the findings to a small set of intuitive or straightforward parameters. It is impossible for the range of assessment of psychometric test to take into account of the full breadth of immersive parameters since immersion is still a lesser-known phenomenon. To identify more surprising or unexpected elements of immersion, perhaps we need firstly to resort to qualitative methods, such as open- or semi-structured interviews, then incorporate these newly-found elements into the questionnaire, to probe their statistical validity.
- iii) Since not all parameters of immersion carry equal weight in contributing to the overall immersive experience, the sensitivity of these psychometric items might be put in question. This is due partly to the nature of immersion that certainly arouses different levels of somatic experience and entails fluctuations in the psychological processes, partly to the fact that immersion may happen in different system, contextual and user conditions. Although more significant influencing factors might be identified through principal component analysis (PCA), minor yet no less significant factors might be neglected or discarded in the analysis. This may very possibly distort the truth, causing considerable bias in evaluating immersive experience.
- iv) Psychometric tests return categorical values, which are limited in representing the breadth and depth of human perception. Even if they probe into the multi-dimensional characteristics of immersion with considerable question items, the understanding with each single question statement is still multi-faceted and varied in dimensions, and cannot be satisfactorily represented along the linear ordinal values. And however these question items are reduced and broken into smaller aspects, the multi-dimensional characteristics of each single question item cannot

be exhausted, since human perception is in itself always multi-dimensional and varies in great breadth and depth. Thus psychometric tests run shorthand of providing information with great subtleties and fail to acknowledge the richness of human perception. However, this can on the other hand be argued as somehow simplifying the experimental design and interpretation by cutting through the clutter and getting down to the nitty-gritty of our essential understanding of immersion.

- v) Psychometric questionnaires usually require the users to respond to a scale somewhere between 3 to 11 points. Thus, not only does the number of points on the scale need to be carefully designed to accommodate the fine-grained human perception, but also this is based on an ungrounded hypothesis that human perception can be satisfactorily represented by being evenly distributed and graded along a linear model. This is a premise that is neither justified by scientific research nor considered realistic in the real-life settings.

2) CONTINUOUS SUBJECTIVE MEASURES

Traditional QoE assessment methods for rating subjective quality of audiovisual content such as the continuous rating methods are not very suitable for measuring immersive experience. Since these methods require the user to manage a slider or rater to continuously judge the multimedia content that has a quality curve during broadcasting, yet managing a slider or rater or even being mindful of the fact that they are immersed is in itself a disruptive activity to the state of immersion. In addition to its intrusiveness, from the continuous measurement we can only obtain uni-variant data – either the overall immersiveness or one aspect of immersion is rated – thus the many-facet richness of immersive experience is undermined. Thus these continuous rating methods are not recommended for measuring immersive experience. Despite these, the advantages of continuous measurement can be: 1) It eliminates the recall error and anchoring effect; 2) It captures the fluctuations of immersive states on a time-variant basis, thus the sensitivity of the results is assured; and 3) It resembles the continuous measurement in QoE, thus a whole set of standardized, mature and rigorous experiment paradigms can be referenced.

B. OBJECTIVE TESTS

1) PRIMARY OR SECONDARY TASK PERFORMANCE

The primary or secondary task performance method is asking the participant to respond to a disturbing secondary task that requires a certain degree of cognitive or attentional allocation during the primary task where the participant is initially immersed in performing. The secondary task can come in many forms, from as simple as answering a “Beep!” from the system sound, to as difficult as those that require some motor and proprioceptive abilities. The secondary task does not necessarily have to be implemented only once, and there

can be intervals of regular or random length between two secondary tasks, depending on what you want to measure. Reaction time or error rates or task completion time from the secondary task are some of the most important objective metrics where data are collected.

This method is based on the theory that immersive experience is a multidimensional construct where the continuous increment of mental workload (i.e. cognitive load) can be measured as an indicator of the extent to which the user's mental capacity and attentional resources are absorbed into the task. This means the more one's attention is allocated to the mediated virtual environment, the greater degree an immersion is indicated, thus the greater portions of external stimuli from the physical reality are ignored, and the more difficult for one to switch his attention back to the real environment, where immediate responses to a secondary task are solicited.

This method has several advantages: i) Since the independent variables it measures fall into the realm of measuring "attention", it can easily combine, within the same experiment protocol, with other attention-measuring methods such as eye-tracking, to obtain a rich variety of data formats for analysis. ii) It is very suitable for measuring the experiential qualities of digital artefacts associated with immersion since its dependent variable is task performance and task performance is directly related to user interacting with the system thus has many experiential dimensions. For instance, "fluency" is an experiential quality which refers to "the degree of gracefulness with which the users deal with multiple demands for their attention and action." Reference [29] The relationship between "fluency" and "immersion" can be properly measured with secondary or triple task performance. This method is also a fitting apparatus to measure other experiential qualities of digital artefacts, if the experiment protocols are thoughtfully designed, such as "transparency" (the degree of opaque a digital artifact is displayed in the design space [30]), "pliability" (the degree to which interaction feels involving, malleable, and tightly coupled, and hence to what degree it facilitates exploration and serendipity in use [31]), or "rhythm" (the human propensity towards rhythmical patterns and temporal predictability. According to Löwgren [32], a certain hypnotic and addictive pleasure may be found in the rhythmic repetition of a motor sequence in a micro-automatic fashion without breakdowns), etc. The correlations of these experiential qualities with immersion are critical to understanding the experiential dimensions of immersion – the aspect of how human users appropriate digital devices and develop an emotional attachment with them. iii) Unlike the psychometric tests which are used as measuring tasks of similar nature to compare immersiveness or usability across devices, contents, or applications, the primary tasks and the secondary tasks in this task performance measurement can be of completely different nature, such as switching from a task of purely cognitive nature to a motor activity. Thus the convergent and discriminant validities of the measurement are ensured. And this gives a lot of space

to exercise flexibility and creativity from the perspective of designing the experiment, and the novelty of the research design also lies in this.

This method also has some flaws in that: i) The secondary task is intrusive to the immersive experience from the primary task, and this might, to a large extent, deflect the face validity of the measurement. ii) The research design can be rather complex and this may involve considerable confounding factors and serendipity effect, thus the reproducibility or test-retest validity can be in doubt. iii) It is difficult to evaluate one's task performance on tasks of significantly different nature, and this may sometimes run the risks of yielding misleading results, thus the internal consistency reliability of the measurement is also questionable. iv) Finally, and most importantly, human cognitive load can be influenced by many factors [33], [34] which are not strictly limited to the immersiveness of the system. Thus, using mental workload theory to test an immersive experience has to consider and evaluate the impacts of certain confounding factors, particularly those from the user and contextual aspects.

2) NEURO-PSYCHO-PHYSIOLOGICAL METHODS

Considering the fact that the biological nature of immersion is a series of neuro-psycho-physiological processes activated or aroused in the users, it would be more effective to directly measure these processes while user is interacting with XR to gain in-depth insights to the mechanisms of immersion.

Given the diversities of these neuro-psycho-physiological processes, the measurement paradigms of them are different. Those metrics measured can be, but are not limited to [2]: electrocardiogram (ECG) from cardiovascular measures such as heart rate and blood pressure, respiratory sinus arrhythmia (RSA) as a measure of controlled attention, phasic heart rate deceleration as a measure of automatic attention, electrodermal activity (EDA) or galvanic skin response (GSR) as measures of skin conductance, amplitude of saccades and scanpath length as measures of spatial eye-tracking, fixation and scanpath duration as measures of temporal eye-tracking, pupillometry (i.e. dilation and contraction of the pupils) as measures of pupil response, facial electromyography (EMG) as a measure of facial expression and emotion, electroencephalogram (EEG) and functional magnetic resonance image (fMRI) as measures of mental activities such as those of a cognitive and emotional nature. The benefits of these measures lie in the fact that they collect data from the participant's automatic and involuntary bodily responses, which are unwavering by the participant's subjective interpretation.

The underlying rationale of this measurement is the response similarity theory [6], assuming that "as the fidelity of the displayed environment increases, responses to that environment will be increasingly similar to responses we exhibit to the same objects, agents, or events in real environments." The more we are orienting ourselves towards the virtual environment, the more we deem it as real and think and behave like it were real.

The advantages of this method include:

- 1) Compared with other methods, it is an accurate measurement that directly reflect the nature of immersion. Since these neuro-psycho-physiological data are rather consistent with one another in their capacities to indicate immersion, they can yield convincing results with almost any target stimuli regardless of the broadcasting device or application. It can also discriminate very well between stressful and relaxing stimuli, stimuli that arouse approach motivation or avoidance motivation, stimuli with or without haptic feedback, and stimuli with considerable surprise and novelty or merely mellowly pleasant. With proper data analysis, it can reveal hidden traits of immersion that previously went unnoticed by other methods, such as those from the psychometric tests.
- 2) This method is non-intrusive to the immersive experience. The electrodes or other wearable devices are very light weighted to the extent that their impact on our biological processes is minimized.
- 3) It captures the participant's objective data thus it is not wavered by the participant's subjective opinions or personal judgments, thus is less prone to errors that are incurred by subjective accounts.
- 4) It captures continuous and time-variant data from which we know exactly when immersions happen, and this gives us the hints to design better immersive experiences by going back and forth to precisely calibrate and re-configure the system factors. For instance, for a storytelling content, if the peak values of the neuro-psycho-physiological signals map the climax in the story development, we can confirm that an immersion happens. In this regard, this method can be used for precisely testing the rhythm and intensity of the plot, making the script-writing more rhythmic and attentive to the step-by-step emotional building-up of crescendos towards the climax.
- 5) Unlike other methods which are measuring user interacting with the system where the human-computer interaction effect can be prominent, this method measures purely the user's biological processes, and by this we estimate the immersiveness of the system. With this understanding the XR system is a transparent and natural extension of our body. And this is exactly what is defined as an "Extended Reality." Thus the neuro-psycho-physiological method is one that measures a reality that is most closely reflecting the realities in XR. And this huge emphasis on the user aspect of immersion accentuates the importance of human factors, highlighting the experiential qualities by systematically quantifying them, and inspiring the human-centered design of immersive experience.
- 6) Since the adverse effects of virtual reality such as cyber-sickness or fatigue are proved to be an easily differentiable and distinct neuro-physiological phenomenon

with biological signals, it would come in handy to detect, measure, and evaluate these phenomena with the neuro-psycho-physiological methods, to devise mitigations to these adverse effects.

The disadvantages of this method include:

- 1) Given the current understanding, these neuro-psycho-physiological processes do not have a one-on-one mapping with the nuanced dimensions of immersion, e.g. different stimuli may arouse similar biological signals, and different signals may point to a similar level of immersion. It is difficult to interpret these signals that allow for a clear indication of what process exactly they measure. Thus the face validity of this measurement can be low.
- 2) Neuro-psycho-physiological signals are very sensitive, to the extent that any noise caused by confounding factors may also be duly recorded. For instance, this method is largely influenced by user's personal idiosyncrasies such as their sudden awareness of being online with the experiment, and these may arouse disturbing emotional, physiological and/or neurological changes that are irrelevant to the test stimuli and thus may greatly distort the results or even overwhelm the desired effect. Pre-screening for mentally healthy participants using neurological methods can help, but cannot completely solve this issue. Yet this can be mitigated by statistical analysis of a large sample group. Still, individual differences in the baseline physiological levels can also interfere with the outcome, thus establishing a baseline for each participant before measuring their immersive experience shall be considered and this facilitates data interpretation.
- 3) The evaluation of biological signals has to take into account of the significant different regulatory focus of spatial immersion and emotional immersion. Spatial immersion is inclined to prevention regulatory focus with avoidance motivation. This means for spatial stimuli, the more immersed the user, the more calm and relaxed they are, and thus the more smooth the signals. And emotional immersion triggers promotion regulatory focus with approach motivation. This means for narrative stimuli, usually the more immersed the user, the more aroused and intensified they are in emotions, and thus the more bumpy the signals. Thus the immersiveness of spatial stimuli and narrative stimuli has to be evaluated differently, otherwise it will yield confusing or even misleading understanding.
- 4) The benefits obtained from this method may not be commensurable to the complexities associated with it. The cost of purchasing the equipment and training the staff is high, the system set-up requires strict calibration, the data interpretation and analysis entail a steep learning curve, and even worse, the participants may face a certain degree of health risks due to potential system failure.

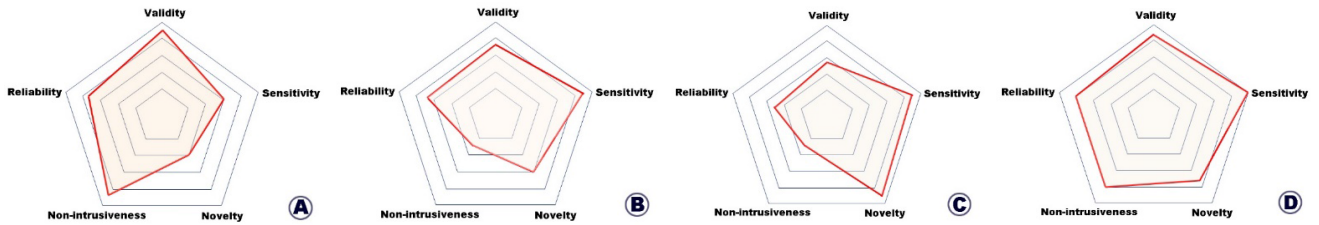


FIGURE 1. Comparing the four measurements along five quality dimensions A – Psychometric questionnaires; B – Continuous subjective measures; C – Primary or secondary task performance; D – Neuro-psycho-physiological method.

V. DISCUSSIONS AND CONCLUSION

Understanding immersion is a difficult task, albeit an important and urgent one. Measuring immersive experience can be used as a unique tool to understanding human cognition, emotion, and many other neuro-psychological processes. The rapid development in reality technologies allows for no delay in the developments of both the theories that grasp the grips of immersion and its practical measurements. This paper has made an initial attempt in this direction. As a global view, we provide a summary evaluation for the four measurements, based on five dimensions of validity, reliability, sensitivity, non-intrusiveness, and novelty. Apparently, the ones with the larger covering areas are recommended (See Figure 1). This figure serves as a natural conclusion by visualizing what we've discussed throughout the paper.

Going back to the four aims and criteria of measuring the quality of immersive experience proposed at the beginning of Section IV, we now can answer them confidently and concretely:

- 1) To measure the “perceptual authenticity” of an XR system, we need to know what makes it perceptually authentic. An important property is its “perceived realness” and a high fidelity interface blurs the fine line between the real and the virtual environment. To know if an XR system appears real, based on the response similarity theory, we need to measure if the user would respond similarly to a scenario in the virtual environment as if they would respond it in the real environment. Then the solution comes: we could establish a baseline to test the user's response in a real environment. The stimuli in the real environment mimicks and approximates the scenarios that would happen in the virtual environment. This is kind of a reverse engineering – if the virtual looks like the real, then the real should look in the same way as the virtual. Then a whole repertoire of neuro-psycho-physiological data from the user's responses to the real events could be recorded. Finally we go back to measure the user's responses to the virtual artefacts. By comparison of these two sets of data, we know if an XR system is “perceptual authentic” and thus immersive. To not to complicate the matter, we could start from measuring responses to relatively simple events or artefacts, then progress to measure the responses to more complicated scenarios.

- 2) To measure how the virtual environment submerges the perceptual system of the users and engulfs their senses, we could start from measuring how and to what extent the XR system blocks the affordances of the perceptual and sensory systems of the users from the real environment. Affordance means what the senses and perceptions are designed to be used for, for instance, the touch system (e.g. the hands) is afforded primarily to grasp objects, the olfactory system (e.g. the nose) is afforded to smell scents, and the auditory system (e.g. the ears) is afforded to listening to sounds, etc. Knowing this, we could design secondary tasks that provide facilities to these affordances. If the user is immersed in the primary task, then they would have difficulty in realizing the affordances in the second tasks, because their senses and perceptions to the stimuli in the real environment are superseded by their immersive experience in the virtual environment.
- 3) All these methods could very well measure attention. Then what about combining them into an integrated measurement of attention? There are existing methods that integrate secondary task performance with eye-tracking in a single experiment, then we could design a method that add psychometric tests to that experiment, too. We could insert the questions from the psychometric tests that are related to probing attention as secondary tasks. In this way we could measure both the time span to cognitively switching from the primary task to the secondary tasks, and at the same time explore the time-variant potentials of the psychometric tests. A caution to this method is that the questions or statements per se that probe into the attentional process should be carefully worded in plain language so that they are easy to understand, to the extent that there is no cognitive impasse in taking them literally. Otherwise the error rates from the secondary tasks would be converted to the validity and reliability issues in the psychometric tests. In addition, some argue that the mental shifts in dual task performance could result in sudden neuro-psycho-physiological changes in the users, thus continuous measurement of these biological signals in dual task performance could detect and record their state of immersion during the switches between tasks [35], [36], by which sifting through the background noise which distinct biological signals may actually map with the

truly immersive state. In particular, users may have the most immersive state(s) and the least immersive state(s) during the process of immersion, especially considering that human beings have attention spans. Using neuro-psycho-physiological measures could identify these extreme moments, and break-in-presence (i.e. secondary tasks) could be inserted during these very moments. And then neuro-psycho-physiological measures could continuously document the biological signals during these abrupt changes. These would help us understand the human cognitive process during immersion, by which revealing the nature of immersion.

- 4) Since immersion is about spatial and emotional perspective taking, using primary or secondary task performance can test the immersiveness of them. For instance, we can alternate spatial tasks and emotional tasks and use reaction time to each task to measure their degree of immersiveness. In this way we could both compare and measure the immersiveness of spatial immersion and emotional immersion, and at the same gain an understanding of the interaction effect of them by examining the time-variant nature of the reaction time. According to Liebold et al. [36], sudden mental violation may lead to inconsistency of mental models in need of compensation or Gestalt. Thus, investigating the dual task performances alternating between spatial immersion and emotional immersion could help reveal the cognitive mechanisms underlying these two fundamental types of immersion.

By discussing these measurements, we aim to enrich the instruments and techniques of measuring immersion, and provide creative sparks that inspire new ideas to either improve the existing methods, or mix and match them to produce hybrid methods, or design entirely new methods based on our introduction. In this paper we discussed the measurement of immersion in controlled laboratory settings. Perhaps immersion could be better measured while in the wild? We wait to see the new test methods and experiment paradigms in this direction. In addition, users might have ad hoc psycho-physiological upsurges and fluctuations when immersed in the virtual system, but the medium- and long-term impacts of the immersive experience that are brought to the users, particularly at the cognitive, emotional and behavioral levels, if any, have not been fully assessed. Knowing these will be particularly conducive to the domain of applying immersive technologies as a therapeutic tool, such as mental or physical rehabilitation. How these long-term post-intervention effects of immersive experience can be traced and measured, in both the laboratory and naturalistic settings, is a new research question that lies before us.

REFERENCES

- [1] C. Zhang, A. S. Hoel, and A. Perks, "Quality of alternate reality experience and its QoE influencing factors," in *Proc. 2nd Int. Workshop Multimedia Alternate Realities (AltMM)*, Oct. 2017, pp. 3–8.

- [2] J. Van Baren and W. IJsselstein, "Measuring presence: A guide to current measurement approaches," Eur. Commission, Brussels, Belgium, Deliverable 5 of the OmniPres Project IST-2001-39237, 2004.
- [3] C. Zhang, A. Perks, and S. Arndt, "Spatial immersion versus emotional immersion, which is more immersive?" in *Proc. 9th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, May/Jun. 2017, pp. 1–6.
- [4] D. P. Salgado, F. R. Martins, T. B. Rodrigues, C. Keighrey, R. Flynn, E. L. M. Naves, and N. Murray, "A QoE assessment method based on EDA, heart rate and EEG of a virtual reality assistive technology system," in *Proc. 9th ACM Multimedia Syst. Conf.*, Jun. 2018, pp. 517–520.
- [5] T. Baumgartner, L. Valko, M. Esslen, and L. Jäncke, "Neural correlate of spatial presence in an arousing and noninteractive virtual reality: An EEG and psychophysiology study," *CyberPsychol. Behav.*, vol. 9, no. 1, pp. 30–45, Feb. 2006.
- [6] W. IJsselstein, "Presence in depth," Ph.D. dissertation, Dept. Ind. Eng. Innov. Sci., Eindhoven Univ. Technol., Eindhoven, The Netherlands, 2004.
- [7] S. Möller, J.-N. Antons, J. Beyer, S. Egger, E. N. Castellar, L. Skorin-Kapov, and M. Sužnjević, "Towards a new ITU-T recommendation for subjective methods evaluating gaming QoE," in *Proc. 7th Int. Workshop Qual. Multimedia Exper. (QoMEX)*, May 2015, pp. 1–6.
- [8] C. C. Bracken, G. Pettey, and M. Wu, "Revisiting the use of secondary task reaction time measures in telepresence research: Exploring the role of immersion and attention," *AI Soc.*, vol. 29, no. 4, pp. 533–538, Nov. 2014.
- [9] C. Klimmt, T. Hartmann, A. Gysbers, and P. Vorderer, "The value of reaction-time measures in presence research: Empirical findings and future perspectives," in *Proc. 8th Int. Workshop Presence*, 2005, pp. 293–298.
- [10] M. Batty, "Digital twins," *Environ. Planning B, Urban Anal. City Sci.*, vol. 45, pp. 817–820, Sep. 2018.
- [11] J. M. Loomis, "Distal attribution and presence," *Presence, Teleoperators Virtual Environ.*, vol. 1, no. 1, pp. 113–119, Jan. 1992.
- [12] C. Sas and G. M. P. O'Hare, "Presence equation: An investigation into cognitive factors underlying presence," *Presence, Teleoperators Virtual Environ.*, vol. 12, no. 5, pp. 523–537, Oct. 2003.
- [13] F. Biocca and B. Delaney, "Immersive virtual reality technology," in *Communication in the Age of Virtual Reality*, F. Biocca and M. R. Levy, Eds. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1995, pp. 57–124.
- [14] M. T. Palmer, "Interpersonal communication and virtual reality: Mediating interpersonal relationships," in *Communication in the Age of Virtual Reality*, F. Biocca and M. R. Levy, Eds. Hillsdale, NJ, USA: Lawrence Erlbaum, 1995, pp. 277–302.
- [15] S. Turner, C. W. Huang, L. Burrows, and P. Turner, "Make-believing virtual realities," in *Digital Make-Believe*. Cham, Switzerland: Springer, 2016, pp. 27–47.
- [16] M. Slater and S. Wilbur, "A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments," *Presence, Teleoperators Virtual Environ.*, vol. 6, no. 6, pp. 603–616, Dec. 1997.
- [17] J. Löwgren, "Fluency as an experiential quality in augmented spaces," *Int. J. Des.*, vol. 1, no. 3, pp. 1–10, 2007.
- [18] C. C. Bracken, G. Pettey, T. Guha, and B. E. Rubenking, "Sounding out small screens and telepresence: The impact of audio, screen size, and pace," *J. Media Psychol. Theor. Methods Appl.*, vol. 22, no. 3, pp. 125–137, 2010.
- [19] M. V. Sanchez-Vives and M. Slater, "From presence to consciousness through virtual reality," *Nature Rev. Neurosci.*, vol. 6, no. 4, pp. 332–339, Apr. 2005.
- [20] M. Singer and B. Witmer (1997), "Presence : Where are we now?" in *Design of computing systems, Social and ergonomic considerations*, M. Smith, G. Salvendy, and R. Koubek, Eds. Amsterdam, The Netherlands: Elsevier, 1997, pp. 885–888.
- [21] W. A. IJsselstein, "Elements of a multi-level theory of presence: Phenomenology, mental processing and neural correlates," in *Proc. PRESENCE*, 2002, pp. 245–259.
- [22] M. Slater and A. Steed, "A virtual presence counter," *Presence, Teleoperators Virtual Environ.*, vol. 9, no. 5, pp. 413–434, Oct. 2000.
- [23] J. V. Draper, D. B. Kaber, and J. M. Usher, "Speculations on the value of telepresence," *CyberPsychol. Behav.*, vol. 2, no. 4, pp. 349–362, Aug. 1999.
- [24] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence, Teleoperators Virtual Environ.*, vol. 7, no. 3, pp. 225–240, Jun. 1998.

- [25] W. A. IJsselstein and G. Riva, "Being there: The experience of presence in mediated environments," in *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments*, vol. 5, G. Riva, F. Davide, and W. A. IJsselstein, Eds. Amsterdam, The Netherlands: IOS Press, 2003.
- [26] M. T. Jones, "Presence as external versus internal experience: How form, user, style, and content factors produce presence from the inside," in *Proc. 10th Annu. Int. Meeting Presence Workshop*, Barcelona, Spain, 2007, pp. 115–126.
- [27] L. W. Barsalou, "Grounded cognition," *Annu. Rev. Psychol.*, vol. 59, pp. 617–645, Jan 2008.
- [28] J. Murray, *Hamlet on the Holodeck*. Cambridge, MA, USA: MIT, 1997, pp. 68–73.
- [29] J. Löwgren, "Fluency as an experiential quality in augmented spaces," *Int. J. Des.*, vol. 1, no. 3, 2007.
- [30] J. Löwgren and E. Stolterman, *Thoughtful Interaction Design: A Design Perspective on Information Technology*. Cambridge, MA, USA: MIT, 2004.
- [31] J. Löwgren, "Pliability as an experiential quality: Exploring the aesthetics of interaction design," *Artifact*, vol. 1, no. 2, pp. 85–95, Oct. 2007.
- [32] J. Löwgren, "Toward an articulation of interaction esthetics," *New Rev. Hypermedia Multimedia*, vol. 15, no. 2, pp. 129–146, 2009.
- [33] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," in *Proc. 50th Annu. Meeting Hum. Factors Ergonom. Soc.*, vol. 9, 2006, p. 904.
- [34] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds., 1st ed. Amsterdam, The Netherlands: North Holland, 1988, pp. 239–250.
- [35] M. Slater, A. Brogni, and A. Steed, "Physiological responses to breaks in presence: A pilot study," in *Proc. 6th Annu. Int. Workshop Presence*, Aalborg, Denmark, 2003.
- [36] B. Liebold, M. Brill, D. Pietschmann, F. Schwab, and P. Ohler, "Continuous measurement of break in presence: Psychophysiology and orienting responses," *Media Psychol.*, vol. 20, no. 3, pp. 477–501, 2017.



CHENYAN ZHANG received the M.Phil. degree in architecture and the moving image from the University of Cambridge. She had also previously studied and worked with Politecnico di Torino, Lancaster University, and the University of Nottingham. She was a Marie Curie Research Fellow with the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, investigating quality of immersive experience in digital storytelling. She was a recipient of several prestigious funding schemes, including: the British Chevening scholarship, and the Marie Curie research fellowships under the European Commission research and funding programmes FP7 and Horizon2020. She has extensive experiences in urban informatics, human–computer interaction, affective computing, and creative AI.

• • •