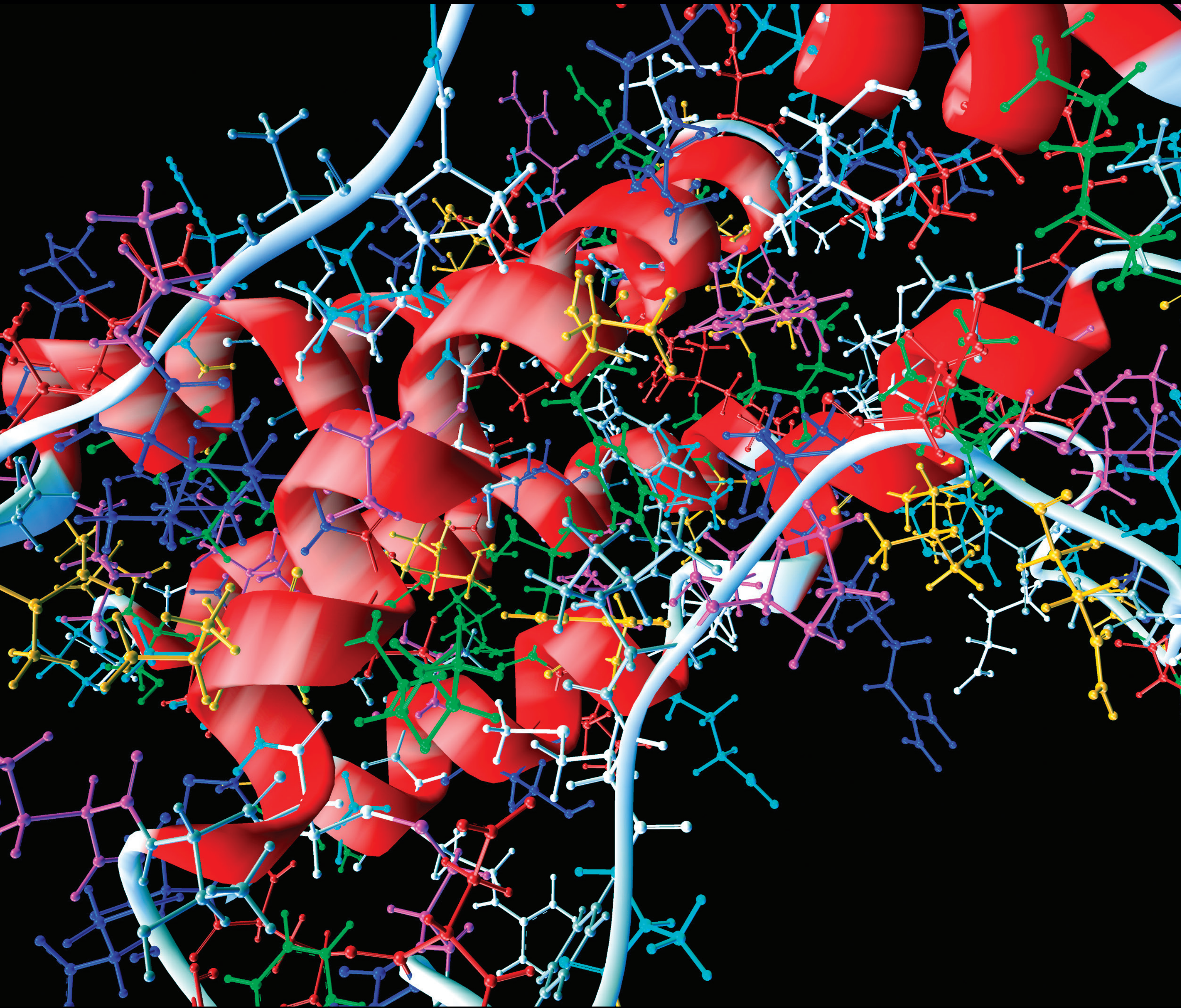
Computational and Mathematical Methods in Medicine

Advances in Computational Psychometrics

Guest Editors: Pietro Cipresso, Aleksandar Matic, Dimitris Giakoumis, and Yuri Ostrovsky



Advances in Computational Psychometrics

Computational and Mathematical Methods in Medicine

Advances in Computational Psychometrics

Guest Editors: Pietro Cipresso, Aleksandar Matic,
Dimitris Giakoumis, and Yuri Ostrovsky

Editorial Board

Contents

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 418683, 2 pages
http://dx.doi.org/10.1155/2015/418683

Editorial

Advances in Computational Psychometrics

Pietro Cipresso,[1] Aleksandar Matic,[2] Dimitris Giakoumis,[3] and Yuri Ostrovsky[4]

[1] *Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Via Ariosto 13, 20145 Milan, Italy*
[2] *Telefonica I&D, Plaça Ernest Lluch I Marti 5, 08019 Barcelona, Spain*
[3] *Information Technologies Institute, Centre for Research and Technology Hellas, 6th Km Charilaou-Thermi Road, P.O. Box 361, 57001 Thessaloniki, Greece*
[4] *Wenzhou Medical College, Wenzhou, Zhejiang 325035, China*

Correspondence should be addressed to Pietro Cipresso; p.cipresso@auxologico.it

Received 6 July 2015; Accepted 6 July 2015

Advances in computational psychometrics and mathematical methods have been gaining a significant role in both medicine and psychology over these past years. The mainstream in psychometrics is moving towards ever greater use of computational and mathematical modeling techniques. Such techniques are critical in the emerging fields of affective and wearable computing, where new biomedical instruments available both in the laboratory and in the field are allowing for deeper understanding of human psychology. These experimental methods offer new opportunities but also new challenges in data interpretation and analysis.

This special issue has two foci, namely, to feature works that (a) advance scientific knowledge in the area of computational psychometrics and (b) explore deep investigated methods, techniques, and instruments for the assessment of cognitive, emotional, and medical (e.g., diagnostic) as well as mental health at the cutting edge of current technology.

There have recently been an increasing number of research initiatives that utilize computational technologies in order to support patients in maintaining or regaining a healthy mental state. Computational psychometrics and related tools have been exploited for assessing, measuring, and defining new methods for an effective and focused psychological intervention.

It is of utmost importance to provide people with higher quality of life and also to shift a part of monitoring tasks from therapists and caregivers to unobtrusive technological systems. Efforts have started with Internet-based self-help therapies, but recently systems make an increasing use of computational psychometrics, including ambient intelligence, pervasive computing, smart phones, and sensor systems. Their common goal is to provide effective solutions for maintaining and improving mental health and related assessment.

This special issue received many articles, accepting for publication five exciting contributions to the field.

P. Cipresso et al. describe a promising approach for managing data from the interaction of two communicating individuals by collecting multiple electrophysiological signals and eye movements with computational methods. D. Cardone et al. introduced a promising thermal infrared imaging-based tool for the computational assessment of human autonomic nervous activity and psychophysiological states in a contactless and noninvasive way. I. A. C. Giglioli et al. presented a systematic review in the novel field of augmented reality for the assessment and treatment of psychological disorders, highlighting 13 selected articles after an initial screening of 784 articles emerging from the scientific databases. Last but not least, two articles in the exciting field of virtual reality are published: C. Wilson and A. Soranzo review some current uses for VR environments when examining visual perception and discuss limitations or questions that can arise; and S. Segkouli et al. define a more robust cognitive prediction model, accurately fitted to human data to be used for more reliable interface evaluation through simulation on the basis of virtual models of users with mild cognitive impairment (MCI).

We hope that this special issue will foster wider discussion for these exciting themes in computational psychometrics.

Pietro Cipresso
Aleksandar Matic
Dimitris Giakoumis
Yuri Ostrovsky

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 151702, 7 pages
http://dx.doi.org/10.1155/2015/151702

Review Article

The Use of Virtual Reality in Psychology: A Case Study in Visual Perception

Christopher J. Wilson[1] and Alessandro Soranzo[2]

[1] *Teesside University, Middlesbrough TS1 3BA, UK*
[2] *Sheffield Hallam University, Sheffield S10 2BP, UK*

Correspondence should be addressed to Christopher J. Wilson; christopher.wilson@tees.ac.uk

Received 10 October 2014; Revised 16 January 2015; Accepted 17 January 2015

Academic Editor: Pietro Cipresso

Recent proliferation of available virtual reality (VR) tools has seen increased use in psychological research. This is due to a number of advantages afforded over traditional experimental apparatus such as tighter control of the environment and the possibility of creating more ecologically valid stimulus presentation and response protocols. At the same time, higher levels of immersion and visual fidelity afforded by VR do not necessarily evoke presence or elicit a "realistic" psychological response. The current paper reviews some current uses for VR environments in psychological research and discusses some ongoing questions for researchers. Finally, we focus on the area of visual perception, where both the advantages and challenges of VR are particularly salient.

1. Introduction

The proliferation of available virtual reality (VR) tools has seen increased use in experimental psychology settings over the last twenty years [1–4]. For the researcher, VR is compelling due to the almost limitless possibilities for the creation of stimuli and this has led to spread of VR into domains such as clinical and developmental psychology, which one might not have initially anticipated [5–7]. Once considered to be an "answer without a question," VR is now firmly established as an experimental tool [8]. However, in addition to the many advantages associated with the use of VR, there remain some drawbacks and ongoing questions. Of course, the relative importance of these issues is dependent entirely on the use case; while presence may be important in a clinical setting, for example, issues with space perception may limit the accuracy of a physical reach task. Similarly, in the experimental examination of visual perception, potential differences between actual and virtual reality can either be advantageous or detrimental. In this paper we provide a brief overview of the benefits and challenges associated with VR in psychology research and discuss its utility in relation to the examination of visual perception.

The term VR is often used interchangeably to refer to one of three types of system: a virtual environment presented on a flat screen, a room-based system such as a CAVE, or a head-mounted display (HMD: [9, 10]). Though all three systems are quite different, a common feature of all three is the introduction of stereoscopic depth, which creates the illusion that the viewer is seeing objects in a virtual space [11]. This offers a number of immediate advantages to the researcher: greater control over stimulus presentation, variety in response options, and potentially increased ecological validity [12]. This has led to increased use of VR as a research tool across many psychological domains such as psychotherapy [13, 14], sports psychology [15], and social interaction [16].

The most apparent advantage of VR is the ability to present stimuli in three dimensions. This offers specific benefits depending on the research domain. For example, when discussing the potential application of VR to neuropsychological research, Rizzo et al. [17] describe virtual environments as "the ultimate Skinner box," able to present a range of complex stimulus conditions that would not be easily controllable in the real world and enabling the examination of both cognitive processes (e.g., attention) and functional behaviours (e.g., planning and initiating a series of required actions). In clinical research VR is used to create complex scenarios, such as simulating exposure to a phobic stimulus, where the form and frequency of the exposure can be manipulated with absolute precision [4]. These examples highlight the difference between VR stimulus presentation and traditional experimental procedures: in VR the participant responds to pertinent stimuli while immersed in a larger virtual environment which can itself be controlled. This differs from traditional experimental contexts where the pertinent stimuli may be controlled but the surrounding environment often cannot be.

Of course, if VR were only a visual medium, then it could be argued that its only advantage over traditional experimental protocols is the ability to present visual stimuli along a third dimensional plane. However, as VR technology has advanced, many VR research studies now include varying levels and combinations of multimodal sensory input, allowing audio, haptic, olfactory, and motion to be experienced simultaneously to the graphically rendered environment or objects [18–20]. This greatly increases the user's sense of immersion in the virtual environment and allows the experimenter to create protocols that would not otherwise be possible. For example, exposure therapy is a common method employed in the treatment of anxiety disorders which, in the case of PTSD, may be difficult to implement for logistical or safety reasons. To overcome these issues, multimodal VR has been employed to create a virtual replica of a warzone, complete with audio and haptic feedback, to treat PTSD in war veterans [21, 22]. Where phenomena are known to occur due to a confluence of sensory data (e.g., audio and visual), multimodal VR enables the researcher to manipulate each input separately to gain a more accurate understanding of the relative contribution of each. For example, a recent study by Keshavarz et al. [23] employed this technique to assess the effects of auditory and visual cues on the perception of vection and resultant motion sickness in participants. Finally, multimodal environments are associated with faster mental processing times of discrete stimulus events, potentially because they provide the user with more complete information about the environment [24].

In addition to the presentation of experimental stimuli, VR enabled researchers to develop new protocols to measure participant responding. Many researchers have no doubt lamented the situation where studies that aim to assess a complex psychological construct (e.g., attention) have, out of necessity, been reduced to a mere "point and click" exercise for the participant. Most experiments strike a difficult balance between control and ecological validity, and very few replicate the multifaceted nature of real-life human responding [25]. It has been suggested that VR environments might help bridge this gap by allowing participants to respond in a manner that is more natural [26]. This can be seen across a range of psychological topics. For example, studies on altruism or prosocial behaviour are often carried out using hypothetical scenarios and self-report responses [27]. Kozlov and Johansen [2] on the other hand, employed a novel approach to examining this topic using VR. As participants attempted to navigate out of a virtual maze, under time pressure, virtual avatars approached the participant for help in a variety of situations. This enabled the experimenters to measure actual helping behaviours, as opposed to participants reporting what they would hypothetically do in such a situation. The researchers argue that even sophisticated high-level behaviours can be successfully examined using VR and suggest wider adoption. VR environments have also been used recently to examine the avoidance behaviour, a central component of fear that contributes to the maintenance of anxiety disorders. While many studies have examined the physiological and self-report aspects of fear, few have been able to examine the associated avoidance of, for example, the context or environment that elicits the fear response [28]. Glotzbach et al. [29] were able to directly examine avoidance behaviour by conditioning participants to be afraid of particular virtual environments and recording the extent to which they avoided returning to those environments later in the study. Finally, VR could be useful to measure responses in circumstances where it might be impractical or ethically questionable to do so in real life. For example, Renaud et al. [30] used a virtual environment and avatars to examine sexual affordances of convicted child molesters. The VR setup allowed the researchers to identify specific patterns of gaze behaviour exhibited by the experimental and not the control group of participants. They discuss a number of theoretical discussions that emerged from the study by virtue of the "first-person stance" enabled by using VR.

2. Questions about the Use of VR in Psychology Research

Since many of the advantages of VR as an experimental tool are derived from the ability to place the participant inside the scene, it is not surprising that a lot of research has been conducted into the concept of presence—the extent to which the user feels as though they are "really there" [31, 32]. Presence is viewed as crucial to having participants respond the same way in VR as they would in reality but remains a difficult concept to measure objectively [32–34]. Many studies have recorded user's subjective experience of presence and the perceived effect it has on engagement with tasks in a virtual environment [35–38]. Kober and Neuper [39] attempted to measure presence objectively and posit that it is characterised by increased attention toward stimuli in the virtual environment and correspondingly lower attention to VR irrelevant stimuli. They were able to identify distinct ERP patterns associated with increased presence. Furthermore, [32] found differences in the levels of presence elicited by a desktop VR system and a more immersive single-wall VR system, which was characterised by stronger activation of frontal and parietal brain regions, measured using EEG.

One of the determinants of presence is the level of immersion, described as the level of sensory fidelity offered by the VR system [40]. It has many contributing components such as field of view, field of regard, display size, and stereoscopy (not exhaustive) and although many use the terms presence and immersion interchangeably (e.g., [41]), they are very different concepts [31]. Immersion is an objective description of the technical capabilities of the VR system that describes the level of detail with which a virtual environment can be rendered, while presence describes the user's psychological response to said environment. Different users can experience different levels of presence in environment with the same level of immersion, depending on a range of factors such as state of mind. Still, it seems intuitive that a researcher would want higher levels of immersion wherever possible, as a higher-fidelity virtual world would elicit more generalizable responses. Indeed, immersive environments seem to be better remembered by participants [37], elicit more intense emotional responses [42], increase collaboration [43], and more successfully replicate the anxiety associated with real-life stressful situations [44]. At the same time, creating an environment elicits a sense of presence that is not entirely dependent on immersion. Factors such as personality and emotional state also influence presence [45–47]. In a research context, realism might not be determined by visual fidelity but by psychological fidelity: the extent to which stimulus presentation evokes the type of physiological or emotional response one would experience in real life. While immersion might help with this goal, it is not the only determining factor [3].

Indeed, it is not universally accepted that higher immersion is always better with some researchers reporting physical and psychological side effects from exposure to VR. These are collectively referred to as virtual reality-induced side effects (VRISE [48]) and often focus on a general feeling of malaise or perhaps motion sickness experienced by users [49]. The effect was initially believed to be caused by limitations in early VR technologies where there was often a lag between participant movements and the display being updated resulting in a disconnection between the perceptual and motor systems of the user [50]. However, while technological advances have overcome this early limitation, VRISE remain a problem [23, 51, 52]. Although common in most VR users, these side effects vary from person to person and, as such, it is difficult to pin down what aspects of immersion are responsible. While some studies suggest that more immersive HMDs are linked to higher levels of sickness in participants [53], others suggest that there is little difference between the side effects of using standard desktop computer display and a head-mounted VR display [54]. Regardless, it seems that these symptoms are generally mild and quick to subside and there is some evidence that users can adapt with repeated exposure [1, 55–57]. While not all that common in the literature, researchers should also consider potential psychological side effects of VR use, depending on the topic being examined. For example, Aardema et al. [58] found that users who had been exposed to an immersive virtual environment demonstrated increase in dissociative experience including a lessened sense of presence in objective reality as the result of exposure to VR, while Aimé et al. [59] found that VR immersion led to body dissatisfaction amongst users. As VR environments become more realistic and scenarios potentially more complex, another potential confound may arise from what Yee and Bailenson [60] term the Proteus effect, where users in a VR environment change their behaviour depending on how they are represented in the virtual world, though currently this effect seems limited to studies that use third-person view and avatars, as opposed to first person perspective [61].

3. The Use of VR in Visual Perception Research

In many domains, the benefits of VR stem from the ability to create recognisable, three-dimensional facsimiles of real objects in space. As a simplified example, let us imagine a study that asks participants to attend to the environment and respond every time they see a person with a happy face. Here the researcher needs only the object (face) to be presented, to be recognised by the participant, and to measure some level of reaction on the part of the participant. In such a context, the main technical focus in relation to VR is likely to be the visual fidelity of the stimuli—the extent to which the faces can be detailed enough for participants to distinguish their expressions. In the experimental study of visual perception, however, the researcher is concerned with how the stimuli are perceived. Here, VR offers both advantages and drawbacks when compared with real life and traditional experimental apparatus. In the following section, we focus on interesting aspects of immersive VR environments that impact how we examine perception: (1) space and movement and (2) tighter control over the visual scene.

3.1. The Effect of Space and Movement in VR on Perception. One area where complications arise is in the perception of space [62]. Many studies have observed a disparity between judgements of distance and perceptual actions such as reaching [63, 64]. In addition, it has been found that in VR, users consistently underestimate the size of the environment and distance to objects [65]. Although not always replicated (e.g., [66]), this effect has been found to be consistent with binocular and monocular vision [67], with varying field of view [68] and even when providing motion parallax and stereoscopic depth cues to the observer [69]. Bingham et al. [62] provide a useful explanation: what we see in VR as an object is actually a series of images mediated by a display. While the user's vision is focused on the series of images that make up the virtual object, the object itself appears in a different location. As a result, when the user is viewing the object, there is disconnect between accommodation (the fixed viewing distance between the user and the display) and convergence (the user's eyes converging on the virtual object), two processes that are inextricably linked in viewing objects in actual reality. Some studies have suggested that this effect is an issue of perception-action recalibration, while others suggest that walking through the virtual environment with continuous visual feedback is necessary to cause rescaling of the perceived space [70].

On the other hand, there are instances where the disconnection between virtual and actual reality provides opportunities for the examination of perception, which would otherwise not be possible. Mast and Oman [71] used a virtual environment to examine visual reorientation illusions, a phenomenon reported by astronauts where the perceived identity of a surface is changed due to rotation of the entire visual field. This phenomenon is difficult to replicate in real life, as we are surrounded by visual cues in our environment that help us to orient ourselves (e.g., trees grow upwards), as well as the orienting force of gravity, which provides a consistent cue for "down." The authors created an immersive environment (i.e., a room containing various objects) with intentionally ambiguous visual cues so that, due to the placement of objects in the room, it could appear correctly oriented even if the room were rotated by 90°. The researchers were then able to rotate the entire visual scene and examine the effects on perception—something that would be almost impossible to replicate in the physical environment.

In addition to creating new illusions, an immersive environment offers the possibility to examine commonly employed visual illusions in new contexts. Traditionally, illusions to examine perception are designed and employed assuming a stationary point of view and have not been studied thoroughly for a moving observer. By employing an immersive environment it is possible to investigate whether these illusions persist when the observer moves. This would be difficult to carry out using a two-dimensional computer screen setup, due to the fact that the stimuli, and hence the illusion, require the observer to view the screen head on. Bruder et al. [72] introduced the use of VR to investigate how visual motion illusions are perceived for a moving observer. The authors manipulated the optical flow—the change of the light pattern on the observer's eyes when moving in the environment—and found that optic flow manipulation can significantly affect users' self-motion judgments.

Movement can also add ecological validity to the examination of everyday perceptual phenomena. Change blindness, a phenomenon in which changes occurring in a visual scene are not noticed by the observer, occurs in a variety of contexts and its impact is studied in range of applied settings from courtroom eye-witness testimony to driving behaviour [73, 74]. Experimental examinations of the effect are usually done on a computer screen where two similar images are presented one after the other with a short blanking between the two, and observers have to indicate whether the second image is the same as the first one or if a change has occurred [75]. Using VR to create a more ecologically valid examination of the phenomenon, Suma et al. [76] had the observer walk through an immersive virtual environment and found that even large changes in the surrounding environment were unlikely to be noticed.

3.2. Control over the Visual Scene. Virtual reality technology overcomes a number of the limitations of traditional experimental methods by enabling precise control of the spatial distribution of the light in the visual scene as well as distance and position of stimuli. In a real room, it is not possible to manipulate these elements completely independently. However, with virtual reality it is possible to manipulate the distances between the surfaces whilst at the same time maintaining the same photometric relationships (i.e., the amount of light reaching the observers' eyes remained constant). Furthermore, by manipulating objects in three-dimensional space, it is possible to examine the effects of positive and negative parallax which would not be possible using a two-dimensional screen. Moreover, the VR technology allows full control of the amount of light reaching the observers' eyes and of the spatial arrangement of the surfaces in the visual scene.

This level of control is particularly useful when we examine colour perception and particular visual phenomena such as colour contrast phenomenon [77]. The colour contrast phenomenon refers to the condition whereby two surfaces with the same spectral composition are perceived to have a different colour when they are placed against different chromatic backgrounds. It has been shown that this phenomenon depends on perceptual belongingness, the grouping of a set of apparent elements into a perceived whole [78, 79]. As Gilchrist et al. [80] explained, "When the [contrast] display is presented in a textbook, it is perceived to belong to the page of the book and to the table on which the book is lying. Thus, [...] the illusion should be quite weak" (p. 814). Adopting a VR technology prevents surfaces from outside of the experimental display from affecting the experimental examination of the colour contrast phenomenon. Although the contrast phenomenon has been a focus of centuries of debate that has interested scientists and philosophers since Aristotle's time [81], there is still no shared consensus of why it actually happens as some authors attribute its occurrence to high-level factors of the visual process whilst others claim that the phenomenon is due to low-level factors. In an attempt to disentangle these viewpoints, Soranzo et al. [82] studied this phenomenon in VR and provided evidence that the colour contrast phenomenon may be attributed to the summative effect of factors occurring to both high- and low-level factors of the visual process.

4. Conclusions

The proliferation of available virtual reality (VR) tools has seen increased use in experimental psychology settings over the last twenty years. In this review, we outlined the advantages and disadvantages of this technology in psychological research, compared to more traditional apparatus. The advantages of VR are that it allows greater control over stimulus presentation; variety in response options; presentation of stimuli in three dimensions; the creation of complex scenarios; the generation of varying levels and combinations of multimodal sensory input potentially allowing audio, haptic, olfactory, and motion to be experienced simultaneously to the graphically rendered environment or objects; the possibility for participants to respond in a more ecologically valid manner; the precise and independent manipulation of the geometric and photometric relationships between objects; the possibility of examining sophisticated complex participants behaviours, such as avoidance; and the study of situations which can be impractical, dangerous, or ethically questionable to be created in real life.

Additionally, we suggest that although this technology has enormous potential to facilitate new discoveries in psychology, there are certain variables that need to be taken into account by the researcher including the concept of presence—immersion alone is not necessarily sufficient to make the participant feel as if the virtual objects are "really there" and respond accordingly; physical and psychological side effects from exposure to VR (virtual reality-induced side effects). In addition, we considered issues that emerge from use of VR in the examination of visual perception and how comparative differences in the perception of colour, contrast, space, and movement, when compared to real life, can be a concern if the goal is exact replication of perception in the physical world or an advantage when trying to create "impossible scenarios."

Finally, it is worth noting that there are large variations in the size and cost of the various apparatus and in some cases they can be impractical for some settings due to their technological complexities. Until quite recently, the price of immersive HMDs with a good tracker system could be prohibitive. However, HMDs are now becoming cheaper and easier to obtain [83, 84], while virtual reality caves, for example, are still comparatively more expensive and require a large amount of space to install [9, 85]. Nevertheless, VR offers exciting opportunities and we hope to see future work that more thoroughly examines the psychometric properties of this useful research tool.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] L. Gregg and N. Tarrier, "Virtual reality in mental health: a review of the literature," *Social Psychiatry and Psychiatric Epidemiology*, vol. 42, no. 5, pp. 343–354, 2007.

[2] M. D. Kozlov and M. K. Johansen, "Real behavior in virtual environments: psychology experiments in a simple virtual-reality paradigm using video games," *Cyberpsychology, Behavior, and Social Networking*, vol. 13, no. 6, pp. 711–714, 2010.

[3] S. Schnall, C. Hedge, and R. Weaver, "The immersive virtual environment of the digital fulldome: considerations of relevant psychological processes," *International Journal of Human Computer Studies*, vol. 70, no. 8, pp. 561–575, 2012.

[4] S. Scozzari and L. Gamberini, "Virtual reality as a tool for cognitive behavioral therapy: a review," in *Advanced Computational Intelligence Paradigms in Healthcare 6. Virtual Reality in Psychotherapy, Rehabilitation, and Assessment*, S. Brahnam and L. C. Jain, Eds., vol. 337 of *Studies in Computational Intelligence*, pp. 63–108, Springer, Berlin, Germany, 2011.

[5] G. Rajendran, "Virtual environments and autism: a developmental psychopathological approach," *Journal of Computer Assisted Learning*, vol. 29, no. 4, pp. 334–347, 2013.

[6] G. Riva, "Virtual environments in clinical psychology," *Psychotherapy: Theory, Research, Practice, Training*, vol. 40, no. 1-2, pp. 68–76, 2003.

[7] G. Riva, "Virtual reality in psychotherapy: review," *Cyberpsychology & Behavior*, vol. 8, no. 3, pp. 220–240, 2005.

[8] N. Foreman, "Virtual reality in psychology," *Themes in Science and Technology Education*, vol. 2, no. 1-2, pp. 225–252, 2009.

[9] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti, "Surround-screen projection-based virtual reality: the design and implementation of the CAVE," in *Proceedings of the ACM Conference on Computer Graphics (SIGGRAPH '93)*, pp. 135–142, August 1993.

[10] R. M. Taylor II, J. Jerald, C. VanderKnyff et al., "Lessons about virtual environment software systems from 20 Years of VE building," *Presence: Teleoperators and Virtual Environments*, vol. 19, no. 2, pp. 162–178, 2010.

[11] J. P. Wann, S. Rushton, and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," *Vision Research*, vol. 35, no. 19, pp. 2731–2736, 1995.

[12] Y. P. Zinchenko, G. Y. Men'shikova, Y. M. Bayakovsky, A. M. Chernorizov, and A. E. Voiskounsky, "Technologies of virtual reality in the context of world-wide and Russian psychology: methodology, comparison with traditional methods, achievements and perspectives," *Psychology in Russia: State of the Art*, no. 3, pp. 12–45, 2010.

[13] C. Suied, G. Drettakis, O. Warusfel, and I. Viaud-Delmon, "Auditory-visual virtual reality as a diagnostic and therapeutic tool for cynophobia," *Cyberpsychology, Behavior, and Social Networking*, vol. 16, no. 2, pp. 145–152, 2013.

[14] D. Villani, F. Riva, and G. Riva, "New technologies for relaxation: the role of presence," *International Journal of Stress Management*, vol. 14, no. 3, pp. 260–274, 2007.

[15] Y. P. Zinchenko, G. Y. Menshikova, A. M. Chernorizov, and A. E. Voiskounsky, "Technologies of virtual reality in psychology of sport of great advance: theory, practice and perspectives," *Psychology in Russia: State of the art*, vol. 4, no. 1, pp. 129–154, 2011.

[16] P. R. Messinger, E. Stroulia, K. Lyons et al., "Virtual worlds—past, present, and future: new directions in social computing," *Decision Support Systems*, vol. 47, no. 3, pp. 204–228, 2009.

[17] A. A. Rizzo, M. Schultheis, K. A. Kerns, and C. Mateer, "Analysis of assets for virtual reality applications in neuropsychology," *Neuropsychological Rehabilitation*, vol. 14, no. 1-2, pp. 207–239, 2004.

[18] C. J. Bohil, B. Alicea, and F. A. Biocca, "Virtual reality in neuroscience research and therapy," *Nature Reviews Neuroscience*, vol. 12, no. 12, pp. 752–762, 2011.

[19] G. Burdea, P. Richard, and P. Coiffet, "Multimodal virtual reality: input-output devices, system integration, and human factors," *International Journal of Human-Computer Interaction*, vol. 8, no. 1, pp. 5–24, 1996.

[20] D. Navarre, P. Palanque, R. Bastide et al., "A formal description of multimodal interaction techniques for immersive virtual reality applications," in *Human-Computer Interaction—INTERACT 2005*, vol. 3585 of *Lecture Notes in Computer Science*, pp. 170–183, Springer, Berlin, Germany, 2005.

[21] M. Gerardi, B. O. Rothbaum, K. Ressler, M. Heekin, and A. Rizzo, "Virtual reality exposure therapy using a virtual Iraq: case report," *Journal of Traumatic Stress*, vol. 21, no. 2, pp. 209–213, 2008.

[22] R. J. Nelson, "Is virtual reality exposure therapy effective for service members and veterans experiencing combat-related PTSD?" *Traumatology*, vol. 19, no. 3, pp. 171–178, 2013.

[23] B. Keshavarz, L. J. Hettinger, D. Vena, and J. L. Campos, "Combined effects of auditory and visual cues on the perception of vection," *Experimental Brain Research*, vol. 232, no. 3, pp. 827–836, 2014.

[24] D. Hecht, M. Reiner, and G. Halevy, "Multimodal virtual environments: response times, attention, and presence," *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 5, pp. 515–521, 2006.

[25] H. T. Hunt, "Why psychology is/is not traditional science: the self-referential bases of psychological research and theory," *Review of General Psychology*, vol. 9, no. 4, pp. 358–374, 2005.

[26] M. T. Schultheis and A. A. Rizzo, "The application of virtual reality technology in rehabilitation," *Rehabilitation Psychology*, vol. 46, no. 3, pp. 296–311, 2001.

[27] S. M. Garcia, K. Weaver, G. B. Moskowitz, and J. M. Darley, "Crowded minds: the implicit bystander effect," *Journal of Personality and Social Psychology*, vol. 83, no. 4, pp. 843–853, 2002.

[28] C. Grillon, "Associative learning deficits increase symptoms of anxiety in humans," *Biological Psychiatry*, vol. 51, no. 11, pp. 851–858, 2002.

[29] E. Glotzbach, H. Ewald, M. Andreatta, P. Pauli, and A. Mühlberger, "Contextual fear conditioning predicts subsequent avoidance behaviour in a virtual reality environment," *Cognition & Emotion*, vol. 26, no. 7, pp. 1256–1272, 2012.

[30] P. Renaud, S. Chartier, J.-L. Rouleau et al., "Using immersive virtual reality and ecological psychology to probe into child molesters' phenomenology," *Journal of Sexual Aggression*, vol. 19, no. 1, pp. 102–120, 2013.

[31] D. A. Bowman and R. P. McMahan, "Virtual reality: how much immersion is enough?" *Computer*, vol. 40, no. 7, pp. 36–43, 2007.

[32] S. E. Kober, J. Kurzmann, and C. Neuper, "Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: an EEG study," *International Journal of Psychophysiology*, vol. 83, no. 3, pp. 365–374, 2012.

[33] J. V. Draper, D. B. Kaber, and J. M. Usher, "Speculations on the value of telepresence," *CyberPsychology and Behavior*, vol. 2, no. 4, pp. 349–362, 1999.

[34] M. Slater, B. Lotto, M. M. Arnold, and M. V. Sanchez-Vives, "How we experience immersive virtual environments: the concept of presence and its measurement," *Anuario de Psicologia*, vol. 40, no. 2, pp. 193–210, 2009.

[35] E. Giannopoulos, Z. Wang, A. Peer, M. Buss, and M. Slater, "Comparison of people's responses to real and virtual handshakes within a virtual environment," *Brain Research Bulletin*, vol. 85, no. 5, pp. 276–282, 2011.

[36] R. Ma and D. B. Kaber, "Presence, workload and performance effects of synthetic environment design factors," *International Journal of Human Computer Studies*, vol. 64, no. 6, pp. 541–552, 2006.

[37] A. Sutcliffe, B. Gault, and J.-E. Shin, "Presence, memory and interaction in virtual environments," *International Journal of Human Computer Studies*, vol. 62, no. 3, pp. 307–327, 2005.

[38] S. Sylaiou, K. Mania, A. Karoulis, and M. White, "Exploring the relationship between presence and enjoyment in a virtual museum," *International Journal of Human Computer Studies*, vol. 68, no. 5, pp. 243–253, 2010.

[39] S. E. Kober and C. Neuper, "Using auditory event-related EEG potentials to assess presence in virtual reality," *International Journal of Human Computer Studies*, vol. 70, no. 9, pp. 577–587, 2012.

[40] M. Slater, "A note on presence terminology," *Emotion*, vol. 3, pp. 1–5, 2003.

[41] A. M. Grinberg, J. S. Careaga, M. R. Mehl, and M.-F. O'Connor, "Social engagement and user immersion in a socially based virtual world," *Computers in Human Behavior*, vol. 36, pp. 479–486, 2014.

[42] V. T. Visch, E. S. Tan, and D. Molenaar, "The emotional and cognitive effect of immersion in film viewing," *Cognition and Emotion*, vol. 24, no. 8, pp. 1439–1445, 2010.

[43] J.-M. Burkhardt, "Immersion, représentation et coopération : discussion et perspectives de recherches empiriques pour l'ergonomie cognitive de la réalité virtuelle," *Intellectica*, vol. 45, no. 1, pp. 59–87, 2007.

[44] J. H. Kwon, J. Powell, and A. Chalmers, "How level of realism influences anxiety in virtual reality environments for a job interview," *International Journal of Human-Computer Studies*, vol. 71, no. 10, pp. 978–987, 2013.

[45] Y. Ling, H. T. Nefs, W.-P. Brinkman, C. Qu, and I. Heynderickx, "The relationship between individual characteristics and experienced presence," *Computers in Human Behavior*, vol. 29, no. 4, pp. 1519–1530, 2013.

[46] M. Rubin and T. Morrison, "Individual differences in individualism and collectivism predict ratings of virtual cities' liveability and environmental quality," *The Journal of General Psychology*, vol. 141, no. 4, pp. 348–372, 2014.

[47] S. Triberti, C. Repetto, and G. Riva, "Psychological factors influencing the effectiveness of virtual reality-based analgesia: a systematic review," *Cyberpsychology, Behavior, and Social Networking*, vol. 17, no. 6, pp. 335–345, 2014.

[48] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems," *Displays*, vol. 29, no. 2, pp. 58–69, 2008.

[49] A. Murata, "Effects of duration of immersion in a virtual reality environment on postural stability," *International Journal of Human-Computer Interaction*, vol. 17, no. 4, pp. 463–477, 2004.

[50] F. Biocca, "Will simulation sickness slow down the diffusion of virtual environment technology?" *Presence: Teleoperators and Virtual Environments*, vol. 1, no. 3, pp. 334–343, 1992.

[51] P. A. Howarth and S. G. Hodder, "Characteristics of habituation to motion in a virtual environment," *Displays*, vol. 29, no. 2, pp. 117–123, 2008.

[52] N. Sugita, M. Yoshizawa, A. Tanaka et al., "Quantitative evaluation of effects of visually-induced motion sickness based on causal coherence functions between blood pressure and heart rate," *Displays*, vol. 29, no. 2, pp. 167–175, 2008.

[53] J. Moss, J. Scisco, and E. Muth, "Simulator sickness during head mounted display (HMD) of real world video captured scenes," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 52, pp. 1631–1634, 2008.

[54] E. Peli, "The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display," *Vision Research*, vol. 38, no. 13, pp. 2053–2066, 1998.

[55] K. J. Hill and P. A. Howarth, "Habituation to the side effects of immersion in a virtual environment," *Displays*, vol. 21, no. 1, pp. 25–30, 2000.

[56] S. Nichols and H. Patel, "Health and safety implications of virtual reality: a review of empirical evidence," *Applied Ergonomics*, vol. 33, no. 3, pp. 251–271, 2002.

[57] E. C. Regan, "Some evidence of adaptation to immersion in virtual reality," *Displays*, vol. 16, no. 3, pp. 135–139, 1995.

[58] F. Aardema, K. O'Connor, S. Côté, and A. Taillon, "Virtual reality induces dissociation and lowers sense of presence in objective reality," *Cyberpsychology, Behavior, and Social Networking*, vol. 13, no. 4, pp. 429–435, 2010.

[59] A. Aimé, K. Cotton, and S. Bouchard, "Reactivity to VR immersions in women with weight and shape concerns," *Journal of Cyber Therapy and Rehabilitation*, vol. 2, no. 2, pp. 115–126, 2009.

[60] N. Yee and J. Bailenson, "The proteus effect: the effect of transformed self-representation on behavior," *Human Communication Research*, vol. 33, no. 3, pp. 271–290, 2007.

[61] N. Yee, J. N. Bailenson, and N. Ducheneaut, "The proteus effect: implications of transformed digital self-representation on online and offline behavior," *Communication Research*, vol. 36, no. 2, pp. 285–312, 2009.

[62] G. P. Bingham, A. Bradley, M. Bailey, and R. Vinner, "Accommodation, occlusion, and disparity matching are used to guide reaching: a comparison of actual versus virtual environments," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 27, no. 6, pp. 1314–1334, 2001.

[63] S. Aglioti, J. F. X. DeSouza, and M. A. Goodale, "Size-contrast illusions deceive the eye but not the hand," *Current Biology*, vol. 5, no. 6, pp. 679–685, 1995.

[64] J. C. Baird and W. R. Biersdorf, "Quantitative functions for size and distance judgments," *Perception & Psychophysics*, vol. 2, no. 4, pp. 161–166, 1967.

[65] V. Interrante, B. Ries, J. Lindquist, M. Kaeding, and L. Anderson, "Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 2, pp. 176–198, 2008.

[66] J. W. Kelly, A. C. Beall, and J. M. Loomis, "Perception of shared visual space: establishing common ground in real and virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 13, no. 4, pp. 442–450, 2004.

[67] D. R. Melmoth and S. Grant, "Advantages of binocular vision for the control of reaching and grasping," *Experimental Brain Research*, vol. 171, no. 3, pp. 371–388, 2006.

[68] J. M. Knapp and J. M. Loomis, "Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 13, no. 5, pp. 572–577, 2004.

[69] I. V. Piryankova, S. de la Rosa, U. Kloos, H. H. Bülthoff, and B. J. Mohler, "Egocentric distance perception in large screen immersive displays," *Displays*, vol. 34, no. 2, pp. 153–164, 2013.

[70] J. W. Kelly, L. S. Donaldson, L. A. Sjolund, and J. B. Freiberg, "More than just perception-action recalibration: walking through a virtual environment causes rescaling of perceived space," *Attention, Perception, and Psychophysics*, vol. 75, no. 7, pp. 1473–1485, 2013.

[71] F. W. Mast and C. M. Oman, "Top-down processing and visual reorientation illusions in a virtual reality environment," *Swiss Journal of Psychology*, vol. 63, no. 3, pp. 143–149, 2004.

[72] G. Bruder, F. Steinicke, P. Wieland, and M. Lappe, "Tuning self-motion perception in virtual reality with visual illusions," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 7, pp. 1068–1078, 2012.

[73] S. G. Charlton and N. J. Starkey, "Driving on familiar roads: automaticity and inattention blindness," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 19, pp. 121–133, 2013.

[74] R. J. Fitzgerald, C. Oriet, and H. L. Price, "Change blindness and eyewitness identification: effects on accuracy and confidence," *Legal and Criminological Psychology*, 2014.

[75] G. W. McConkie and C. B. Currie, "Visual stability across saccades while viewing complex pictures," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 22, no. 3, pp. 563–581, 1996.

[76] E. A. Suma, S. Clark, S. L. Finkelstein, and Z. Wartell, "Exploiting change blindness to expand walkable space in a virtual environment," in *IEEE Virtual Reality Conference (VR '10)*, pp. 305–306, March 2010.

[77] G. Y. Menshikova, "An investigation of 3D images of the simultaneous-lightnesscontrast illusion using a virtual-reality technique," *Psychology in Russia: State of the Art*, vol. 6, no. 3, pp. 49–59, 2013.

[78] W. Benary, "Beobachtungen zu einem Experiment über Helligkeitskontrast," *Psychologische Forschung*, vol. 5, no. 1, pp. 131–142, 1924.

[79] M. Wertheimer, "Untersuchungen zur Lehre von der Gestalt. II," *Psychologische Forschung*, vol. 4, no. 1, pp. 301–350, 1923.

[80] A. Gilchrist, C. Kossyfidis, F. Bonato et al., "An anchoring theory of lightness perception," *Psychological Review*, vol. 106, no. 4, pp. 795–834, 1999.

[81] N. J. Wade, "Descriptions of visual phenomena from Aristotle to Wheatstone," *Perception*, vol. 25, no. 10, pp. 1137–1175, 1996.

[82] A. Soranzo, J.-L. Lugrin, and C. J. Wilson, "The effects of belongingness on the simultaneous lightness contrast: a virtual reality study," *Vision Research*, vol. 86, pp. 97–106, 2013.

[83] Oculus, *Oculus Rift-Virtual Reality Headset for 3D Gaming*, 2012, https://www.oculus.com/.

[84] Samsung, Samsung Gear V.R., 2014, http://www.samsung.com/global/microsite/gearvr.

[85] N. Firth, "First wave of virtual reality games will let you live the dream," *New Scientist*, vol. 218, no. 2922, pp. 19–20, 2013.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 862942, 12 pages
http://dx.doi.org/10.1155/2015/862942

Review Article

Augmented Reality: A Brand New Challenge for the Assessment and Treatment of Psychological Disorders

Irene Alice Chicchi Giglioli,[1] Federica Pallavicini,[1] Elisa Pedroli,[1] Silvia Serino,[1] and Giuseppe Riva[1,2]

[1] *Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, 20145 Milan, Italy*
[2] *Department of Psychology, Catholic University of Milan, 20123 Milan, Italy*

Correspondence should be addressed to Irene Alice Chicchi Giglioli; alice.chicchi@gmail.com

Received 12 December 2014; Accepted 3 February 2015

Academic Editor: Yuri Ostrovsky

Augmented Reality is a new technological system that allows introducing virtual contents in the real world in order to run in the same representation and, in real time, enhancing the user's sensory perception of reality. From another point of view, Augmented Reality can be defined as a set of techniques and tools that add information to the physical reality. To date, Augmented Reality has been used in many fields, such as medicine, entertainment, maintenance, architecture, education, and cognitive and motor rehabilitation but very few studies and applications of AR exist in clinical psychology. In the treatment of psychological disorders, Augmented Reality has given preliminary evidence to be a useful tool due to its adaptability to the patient needs and therapeutic purposes and interactivity. Another relevant factor is the quality of the user's experience in the Augmented Reality system determined from emotional engagement and sense of presence. This experience could increase the AR ecological validity in the treatment of psychological disorders. This paper reviews the recent studies on the use of Augmented Reality in the evaluation and treatment of psychological disorders, focusing on current uses of this technology and on the specific features that delineate Augmented Reality a new technique useful for psychology.

1. Introduction

Augmented Reality (AR) is a new technological system that allows inserting virtual contents in the real world in order to run in the same representation and, in real time, enhancing the user's sensory perception of reality [1]. Compared to a virtual reality system characterized by a computer-generated environment that elicits a strong user's experience of "presence" [2, 3], an AR system applies virtual and real elements in a real scene augmenting the user's perception of the world.

More precisely, Azuma et al. [4, 5] defined an AR platform as a system that [6]

(i) combines real and virtual objects in a real environment,

(ii) runs interactively and in real time,

(iii) registers real and virtual objects with each other.

Furthermore, according to Milgram et al. [7, 8], AR places between reality (real environment) and virtuality (virtual environment) on the reality-virtuality continuum (see Figure 1).

In an AR system, users see an image made up of a real image and virtual elements that are superimposed over it. The addition of virtual elements may also inhibit the perception of real elements by overimposing the virtual elements on the real elements. Nevertheless, the most important aspect in AR is that the virtual elements provide the real world with remarkable and valuable information. The addiction of virtual elements may involve not only the view but also the hearing, smell, and touch [9].

From the point of view of technology devices, AR can be defined as a set of techniques and tools that allow adding information to the physical reality. Various technologies are used in AR rendering including handheld devices, display

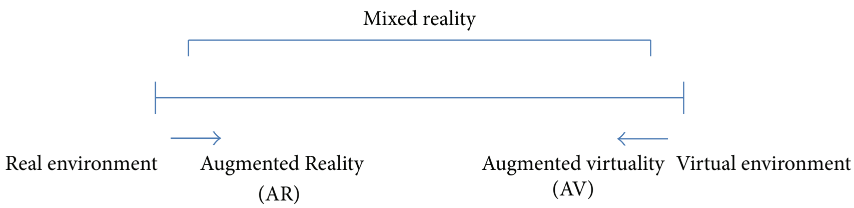


Figure 1: It shows a continuum between reality and virtual reality. Mixed reality is located between them and includes Augmented Reality (AR) and augmented virtuality (AV). AR is placed closer to real enviroment than virtual environment.

system worn on one's user (nonhandheld devices), and projection displays.

Modern handheld mobile computing like smartphones and tablets contain these elements, which include a camera and sensors such as accelerometer, Global Positioning System (GPS), and solid-state compass, making them a suitable AR platform. The smartphone or tablet take in real time the surrounding environment and the virtual elements are superimposed to the real world.

An example of nonhandled device is the head mounted display (HMD), a display system worn on one user's head such as a helmet or glasses. HMD systems are characterized by sensors that receive input information about the environment by the user's head movements adding at a later stage various virtual contents.

As regards the projection displays, the virtual elements are projected on the real objects in order to be augmented. The projection occurs with a single room-mounted projector without any display system worn on one user's head. The projector generates a virtual image on the room surface using an automated calibration procedure that takes into account the structure of the surface overlapping the virtual image.

Furthermore, an AR platform requires a software application able to augment the real world by using one or more hardware devices. Marker-based and markerless systems are the two main software applications used in AR system. The AR marker-based systems are stylized pictures in black and white that are recognized by the computer webcam and which are superimposed in real time multimedia contents: video, audio, 3D objects, and so forth. Instead, in the markerless [10] AR system the software application catches the user's positional and orientation data through GPS and compass device adding the virtual contents in an accurate position on or in the real environment.

To date, AR has been used in medicine [11], entertainment [12], maintenance [13], architecture [14], education [15, 16], and cognitive [17, 18] and motor rehabilitation [19–24] but very few applications of AR exist in clinical psychology and, in particular, it is still underused in the treatment of psychological disorders [25].

Starting from these premises, the aim of this paper is to review the recent studies on the use of AR in the evaluation and treatment of psychological disorders, focusing on current uses of AR in psychology and the various factors that make a new technique useful for the treatment of psychological disorders, expanding the possible fields of use of AR.

2. Materials and Methods

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [26].

2.1. Search Strategy. A computer-based search in several databases was performed for relevant publications describing the use of AR in psychology. Databases used for the search were PsycINFO, PubMed/Medline, and Web of Science (Web of Knowledge). We searched using the string "Augmented Reality" AND (psycholog* OR assessment OR treatment). We excluded articles where the full text was not available or where the abstract lacked basic information for review. The first search was performed for publications in the English language, and then we decided to clean the results, considering only publications for the last ten years, from 2004 forward and eventually updated the search results through December 2014. Expert colleagues in the field were contacted for suggestion on further studies not considered in our search.

2.2. Selection Criteria. We have included articles on AR used for psychological settings in assessment or treatment studies. Excluded from the analysis were studies that omitted the inclusion criteria, non-English published studies, review articles, case reports, letters to the editor, research protocols, patents, and editorials. We tried to contact corresponding authors of the included studies with the intent of obtaining incomplete or supplementary data.

2.3. Quality Assessment and Data Abstraction. To assess a risk of bias, PRISMA recommendations for systematic literature analysis have been strictly followed. Three authors (Irene Alice Chicchi Giglioli, Federica Pallavicini, and Silvia Serino) independently selected paper abstracts and titles and analyzed the full papers that met the inclusion criteria, resolving disagreements through consensus.

The data extracted from each included study were sample type, study design, sample size, type of therapy or assessment, and selected findings.

3. Results and Discussion

3.1. Systematic Review Flow. The flow chart of the systematic review is shown in Figure 2. Our initial search yielded 918 nonduplicate citations screened via PsycINFO, PubMed/

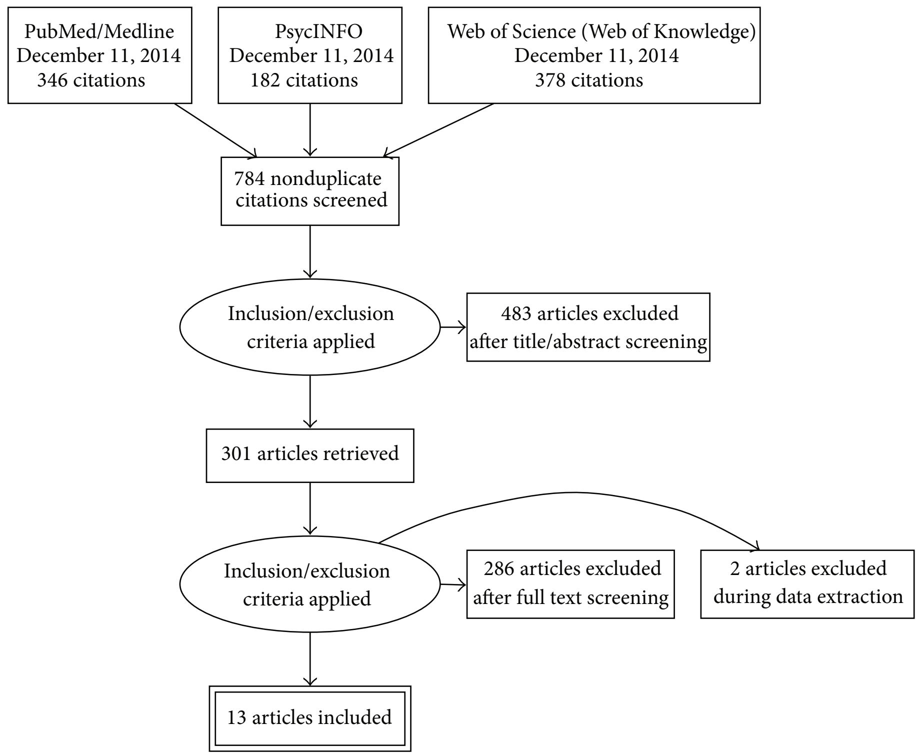


Figure 2: Flow diagram of study selection.

Table 1: Detailed search strategy.

	"Augmented Reality" and psycholog*	Assessment	Treatment	Other sources	Total
Medline	23	56	267		346
PsycINFO	133	21	28		182
Web of Science (Web of Knowledge)	69	145	164		378
Total	225	222	459	12	**918**
Nonduplicated	191	203	378	12	**784**
Excluded (after reading title and abstract)					**483**
Retrieved					**301**
Excluded (after applying inclusion criteria)					**286**
Excluded (missing experimental data)					**2**
Included					**13**

* is a Jolly characters that means that the search strategy included terms as psychology and/or psychological.

Medline, Web of Science (Web of Knowledge), and other sources: more information is available in Search Strategy and supplement. After the application of inclusion/exclusion criteria, papers have been reduced to 301 articles. A more in-depth investigation of the full papers resulted in an exclusion of 286 articles. During the data extraction procedure, 2 additional full papers were excluded. In the end, 13 studies met full criteria and were included in this review (see Table 1).

3.2. Selected Studies on Augmented Reality. Despite a large volume of studies on AR, little has been done specifically to psychological assessment or treatment. In the current review we present a broad range of experimental and clinical studies.

3.3. Results. In the area of clinical psychology, a few but remarkable studies have met the DSM-IV-TR criteria [27], showing the usefulness of AR in the treatment of a specific phobia. More specifically, the review of the literature showed that the phobia of small animals (cockroaches and spiders) and the acrophobia [28–38] are the current areas that used AR in the assessment and treatment of specific psychological disorders. For the readability of the contents, we have divided the results section in two paragraphs based on specific phobia's typology. The studies are presented in chronological order for showing the developments and advancements that occurred in the use of AR in this area. In addition, the selected final studies can be found in Tables 2 and 3. Each of them is described below.

Table 2: Information about the selected studies on the assessment and treatment of specific phobia for small animals using an AR system.

Authors	Year	Sample	Conditions	Dependent variables	AR device	Results
Juan et al. [39]	2004	1 patient with cockroach phobia	Case study with AR	Anxiety	HMD with AR-tags	Decrease of anxiety level after treatment
Botella et al. [35, 40]	2005	1 patient with cockroach phobia	Case study with AR	Avoidance behaviour Degree of belief in catastrophic thought Anxiety Target behaviour Presence Reality judgment	HMD with AR-tags	Decrease of anxiety, fear, and avoidance after treatment High level of presence and judgment of reality
Juan et al. [31]	2005	9 patients with cockroach and spider phobia	AR	Anxiety Fear Avoidance behaviours Presence Reality judgment	HMD with AR-tags	Decrease of anxiety, fear, and avoidance after treatment High level of presence and judgment of reality
Botella et al. [34]	2010	6 patients with cockroach phobia	AR	Anxiety Target behaviour Behavioural avoidance	HMD with AR-tags	Decrease of anxiety, fear, and avoidance after treatment and maintained at follow-up periods (three, six, and twelve months follow-up)
Bretón-López et al. [32]	2010	6 patients with cockroach phobia	AR	Anxiety Presence Reality judgment	HMD with AR-tags	Decrease of anxiety. High levels of presence and reality judgment
Botella et al. [33]	2011	1 patient with cockroach phobia	Case Study with AR	Anxiety Target behaviours Behavioural avoidance	Mobile phone marker less versus HMD AR-tags	Decrease of anxiety, fear, and avoidance behaviours in both conditions.
Juan and Calatrava [37]	2011	24 healthy volunteers	AR-OST HMD AR-VST	Anxiety Presence	AR-OST HMD versus AR-VST	The AR-VST induced grater sense of presence than AR-OST. Significant anxiety in both conditions
Juand and Joele [38]	2011	24 healthy volunteers	AR marker-based versus invisible AR-tags	Anxiety Presence	HMD with visible AR-tags versus invisible AR-tags	The invisible AR-tags induced a higher sense of presence and anxiety than the visible AR-tags
Wrzesien et al. [29]	2011a	22 patients with cockroach and spider phobia	In vivo versus AR	Anxiety, behavioural avoidance Belief in negative thoughts	HMD with AR-tags	Decrease of anxiety, fear, avoidance behaviours, and belief in negative thoughts
Wrzesien et al. [28]	2011b	22 patients with cockroach and spider phobia	AR	Anxiety, behavioural avoidance Belief in negative thoughts	HMD with AR-tags	Decrease of anxiety, fear, avoidance behaviours, and belief in negative thoughts
Wrzesien et al. [30]	2013	26 healthy volunteers	AR	Anxiety Avoidance	Therapeutic lamp (TL)	Decrease of anxiety

3.4. AR and Specific Phobia for Small Animals. In the studies, the AR exposure therapy was based on Öst et al.'s "one-session treatment" guidelines [41, 42]. Individuals make one single intensive exposure session that lasts up to three hours.

The first analysed study that used an AR system to assess and treat specific phobias was conducted by Juan et al. (2004) [39]. A single individual with cockroach phobia [27] was assessed using an HMD-AR system. The AR device used

Table 3: Information about the selected studies on the assessment of acrophobia using an AR system.

Authors	Year	Sample	Conditions	Dependent variables	AR device	Results
Juan et al. [43]	2006	41 healthy volunteers	Real environment versus AR environment	Sense of presence	AR-HMD	High sense of presence in the AR environment
Juan and Prez [36]	2010	20 healthy volunteers	AR system versus VR system	Sense of presence Anxiety	HMD-AR with tags VR-HMD	In regard to sense of presence and anxiety levels, AR is effective as VR

was an HMD system connected with a camera and a PC. The camera, placed on the HMD, recognized the marker through the movement of the subject's head, projecting the virtual cockroaches in front of the subject. The AR single exposure session consisted progressively in seeing, touching, and finally killing one or more virtual cockroaches. The therapist chose in any moment how many cockroaches had to appear on the scene, their size, and if they had to move or not. During the treatment, the augmented cockroaches were able to arouse anxiety in patient that decreased after an hour of exposure. More specifically, before, during, and after treatment, the patient ranked her level of anxiety on a 10-point scale (where 0 represents no anxiety and 10 very high anxiety) using the Subjective Units of Discomfort Scale (SUDS) [44]. The data showed a decrease in anxiety score after exposure (with a score of 10 on SUDS at the beginning of the session and a score of 0 after the session) and clinical improvements regarding patient's phobia. In particular, after exposure, the patient was able to approach, interact, and kill real cockroaches.

Botella et al. (2005) [35] assessed a cockroach phobia case study [27] using an HMD-AR system developed for this specific disorder. The AR device used was an HMD system connected with a camera and a PC. The camera, placed on the HMD, recognized the marker through the movement of the subject's head, projecting the virtual cockroaches in front of the subject. The AR system included the possibility for the therapist to choose the number of cockroaches, their size, and movements and for the patient to kill one or more cockroaches using two different instruments, a fly swatter or a cockroach killer. Depending on the selected tool the system played a sound analogous to the real one. In order to assess the intensity of the phobia, the behaviour avoidance test (BAT) [42], degree of belief in catastrophic thought (assessed daily on scale from 0% to 100%), Fear and Avoidance Scales [45],Fear of Spiders Questionnaire (FSQ) [46], Spider Phobia Beliefs Questionnaire (SPBQ) [47], and Subjective Units of Discomfort Scale (SUDS) [44] were used. Furthermore, in order to assess the degree of presence and reality judgment experienced in the AR session, the authors created specifically for this study one ad hoc questionnaire composed by two questions related to presence: "To what degree have you felt present in the situation?" and "To what degree have you felt that you were in the place where the cockroaches appeared?" and one question for the reality judgment: "To what degree did the cockroaches appear to be real?" After AR exposure, the participant was asked to evaluate these features on a scale from 0 (no degree of being in a place/being real) to 10 (very high degree of being in a place/being real). The results showed that, before the exposure, patient exhibited a considerable fear and avoidance behaviours and after exposure not only were there important decreases in the fear and avoidance scores, but also the patient was able to approach, interact, and kill cockroaches with a high degree of presence and reality judgment. Similarly, at the beginning of the experiment, virtual cockroach induced anxiety in patient but after one hour of exposure the anxiety was significantly declined. Finally, the treatment gains were maintained in a follow-up conducted two months after the end of the treatment, showing decreases in the various scales of the BAT [42].

Juan et al. (2005) [31], for the first time, evaluated the effectiveness of an AR system not in one single case but in a sample of nine patients with cockroach and spider phobia [27]. The AR system was the same used in the Botella et al. study [35], described previously. The AR exposure involved the gradual appearance of one or more spiders/cockroaches and the possibility for the patients to approach them with hands, to look in boxes in order to simulate when you are searching for a small animal in your house, and to beat and throw away them. With respect to psychological measures, anxiety, fear, and avoidance behaviours were assessed using Fear and Avoidance Scales [45] and SUDS [44]. The degree of presence and reality judgment experienced by the users in the AR system were assessed using three ad hoc questions, created specifically for this study. The three questions were "To what degree have you felt present in the situation?" "To what degree have you felt that you were in a place in which spiders/cockroaches appeared?" And "To what degree did you think the spiders/cockroaches were real?" Participants were asked to evaluate these features in a scale from 0 to 10 (where 0 represents the lowest degree of being in a place/being real and 10 the highest degree of being in a place/being real). The results showed that the treatment produced a decrease in the patient's fear and avoidance behaviours [45] when they had to face with the target spider/cockroach. Furthermore, as in the study of Botella et al. [35], during the exposure, the participant's anxiety scores (SUDS) [44] were high, but they diminished at the end of the treatment.

In Botella et al. (2010) [34] an AR system was tested in the short and long term (three-, six-, and twelve-month follow-up) on a sample of six individuals suffering from cockroaches phobia [27]. The AR system was the same as that used in the preceding studies [31, 35]. Before, during, and after the AR exposure, participants were asked to fill out questionnaires, including SUDS [44] to evaluate anxiety levels, BAT [42], FSQ [46], and SPBQ [47] to assess fear and avoidance behaviours. AR exposure has been led in a single extended session lasting up to three hours. Each participant faced with various scenarios, progressing from the easiest to the most difficult situation. For example, at the beginning of the exposure, the program exhibited one cockroach to the participants, and more animals were added progressively. The purpose of the exposure was to interact with many cockroaches repeatedly, touch them, kill them, and remain in the situations until they experienced a considerable decrease in anxiety. As above in Botella et al. (2005), the results showed that AR was effective in treating cockroach phobia, improving significantly in all outcome measures after treatment. More specifically, the AR system, at the beginning of the exposure, was able to induce anxiety in the participants and after treatment produced a significant decrease in the level of fear and avoidance behaviours in all participants. In accordance with the BAT's scores [42], before treatment, none of the participants were able to interact with a real cockroach, while after treatment all participants could approach it. At the same time, also the self-report scores of FSQ [46] and SPBQ [47] improved significantly after treatment. Furthermore, the treatment gains were maintained at three-, six-, and twelve-month follow-up periods. Unlike the study conducted in 2005 by Botella et al. [35], measures related experience has not been recorded.

Bretón-López et al. (2010) [32] evaluated the ability of an AR system to raise anxiety and, secondly, to elicit sense of presence and reality judgment in six participants suffering from cockroach phobia [27]. In the single AR exposure session, participants were exposed to various stimuli, based on each individual's hierarchy of fears (from one static insect to the one in movement, from more static insects to ones in movement, and from insects next to personal belongings to those next to hands). Moreover, the AR system included the possibility to vary the numbers of cockroaches on the scene and the movement and the size of cockroaches. During and after the AR exposure, patients' level of anxiety was evaluated through the SUDS on a scale of 0 (no anxiety) to 10 (very high anxiety) [44], and the degree of presence through the Presence and Reality Judgment Questionnaire (PRJQ) [48]. The data showed that, at the start of AR immersion, the anxiety levels, measured through the SUDS questionnaire [44], ranged from 9 to 10 (the highest levels of the scale) but they decreased progressively during the exposure. The levels and the duration of exposure needed for anxiety reduction were based on initial levels of fear and on the severity of the phobia. Furthermore, patients showed high scores in the PRJQ [48], representing high levels of presence and reality judgment experienced during the AR exposure.

In 2011, Botella et al. [33] assessed a single cockroach phobia case study [27] testing an AR system using a mobile phone and creating a mobile game, "Cockroach Game" [33] for the treatment of this phobia. The subject conducted three therapeutic phases. In the first phase, the participant was asked to play, as much as she wanted, "Cokroach Game" for nine days and to record her levels of fears, avoidance, and belief in the catastrophic thought using the BAT [42] and FSQ [46]. In the second phase, the participant received the AR exposure treatment, assessing again the psychological measures. In the third phase, the participant was asked to play again the "Cockroach Game" for another period of nine days and recorded the same questionnaire. In this study the experiential measures of presence and reality judgment have not been taken into account. The data showed that before and after the first phase a slight improvement is obtained in performance, fear, and avoidance, whereas there was an increased in the belief in the catastrophic thought. After the AR exposure there were improvements in all BAT measures [42]: performance, fear, avoidance, and belief. After the third phase a significant decrease was obtained in all psychological measures, maintaining them at one, three-, six-, and twelve-month follow-up periods.

In 2011, Juan and Calatrava [37] compared an AR optical see-through (OST) system with a video see-through (VST) for the treatment of spiders and cockroaches phobia in twenty-four nonphobic participants. Individuals were divided in low and high fear subject's group according to scores in the fear and avoidance of cockroach and spider questionnaires [46] and underwent both experimental conditions. Before, during, and after exposure subjects were asked to rate their level of anxiety from 0 (no anxiety) to 10 (very high anxiety). Presence experienced by participants was assessed using an adapted version of the Slater et al.'s (1994) questionnaire [49]. The six question were "Please rate your sense of being in a room where there are cockroaches/spiders." "To what extent were there times during the experiment when the cockroaches/spiders were real for you?" "When you think back to your experience, do you think of cockroaches/spiders more as images that you saw (a movie, a picture), or more as cockroaches/spiders that were in the same room as you were?" "During the experiment which was strongest on the whole: your sense of being in the room where there were cockroaches/spiders, or your sense of being in a room without cockroaches/spiders?" "Think about your memory of being in a room where there were cockroaches/spiders. How similar is this memory to your memories of other places where there were these animals?" And "During the experiment, did you often think that you were actually in a room where there were cockroaches/spiders?" The scoring was on a scale of 1–7 (where 0 corresponds to not being in a place and 7 represents the normal experience of being in a place). Results showed that the VST system induced a higher sense of presence than the OST system and the two systems produced similar and significant anxiety before treatment that decreased after exposure.

Juan and Joele (2011) [38] compared an AR visible marker-based system with an AR invisible marker system in twenty-four healthy subjects. Participants rated their intensity of anxiety level from 0 (no anxiety) to 10 (very high anxiety) at eight different moments during the AR exposure. As in the previous study, after each AR exposure, individuals were asked to fill out an adapted Slater et al.'s (1994) questionnaire (SUS) [49] to rate the subjective feelings of presence experienced. Results showed that the AR invisible marker system elicited a higher sense of presence compared to the AR visible marker system. Furthermore, at the beginning of the treatment, the AR invisible marker system provokes a higher level of anxiety that decreases significantly during and at the end of the AR exposure.

Wrzesien et al. (2001a; 2011b) [28, 29] conducted two studies to evaluate the level of anxiety, avoidance, behavioural avoidance, and belief in negative thoughts in patients with small animal phobia [27].

The first study [29] compared the in vivo exposure therapy with the AR exposure therapy in twenty-two individuals with specific phobia for spiders and cockroaches [27]. The patients were randomly allocated to one of two groups. Before and after the exposure session, the participants were asked to fill out the behaviour avoidance test (BAT) [42]. The data showed that both in vivo and AR exposure are therapeutically effective in reducing anxiety, avoidance, and behavioural avoidance. In particular, the analysis of the pre- and posttest BAT scores showed no statistically significant differences between the in vivo group and the AR exposure group. Furthermore, an intragroup analysis showed a statistically significant decline in the severity of avoidance under both conditions, suggesting that both exposures are effective in the reduction of avoidance behaviours after treatment.

In the second study [28], five patients were assessed and treated using only an AR therapeutic exposure [27]. Before, during, and after the exposure session, the participants were asked to rank their level of anxiety on a scale of 0 (no anxiety) to 10 (very high anxiety), avoidance on a scale of 0 (low degree of avoidance) to 10 (high degree of avoidance), behavioural avoidance on a scale of 0 (low degree of behavioural avoidance) to 13 (very high degree of behavioural avoidance), and belief in catastrophic thoughts on a scale of 0 (low degree of belief) to 10 (high degree of belief). The results showed a posttreatment decrease in level of anxiety, avoidance, and belief in negative thoughts. More precisely, if, before the therapy, patients were unable to get close to live cockroaches, after the treatment they were able to interact with real cockroaches into a terrarium.

In a further study, Wrzesien et al. (2013) [30] experimented with an innovative technological AR system named therapeutic lamp (TL). The TL is an AR display projector created for the treatment of small animals' phobia. The trial included twenty-six healthy volunteers and consisted of one single exposure session composed by twelve exercises progressed from those who induced less anxiety to the ones that elicited more anxiety. For example, at the beginning of the exposure, participants observed three dead and three paralyzed animals and, at the end, they had to kill 30 animals with the flyswatter. In order to measure the participants' experience during the exposure, four clinical instruments were used. The Spider and Cockroach Anxiety and Avoidance Questionnaire was assessed before the session on a scale of 0 (no degree of fear and avoidance) to 7 (high degree of fear and avoidance). The Self-Efficacy Belief Questionnaire was used before and after the exposure session on a scale of 0 (no degree of belief that the participant could confront a real cockroach or spider) to 7 (high degree of belief). The Subjective Units of Discomfort Scale (SUDS) [44] was assessed at the beginning and end of each exercise in the session on a scale to 0 (no anxiety) to 10 (very high anxiety). The Presence and Reality Judgment Questionnaire (PRJQ) [48] was tested at each exercise's start on a scale of 0 (no degree of being in a place/being real) to 10 (high degree of being in a place/being real). The data showed that the participants' anxiety scores, measured by SUDS [44], were high at each exercise's beginning but decreased after the exercise session. In addition, the participant's belief in their capacity to face with the small animals improved significantly after the session. Finally, the PRJQ [48] scores showed that the participants felt the virtual animals' presence relatively well and considered them to be rather real. Therefore, the authors concluded that the AR-TL could be a good and helpful treatment's tool for psychological disorders even if the system has to be validated with patients in future studies.

All these studies represent new potentiality and possibility of assessment and treatment in the area of psychological disorders. However, they disclose some limitations. The majority of the disclosed studies [28, 29, 31–35] include a too small sample for the experimental validity characterized by patients with specific phobias. Instead, one study [30] focused on testing an innovative AR system with healthy volunteers in order to verify the efficacy, usability, and quality of user's experience of the new platform. Related to the previous consideration, the presented studies [28, 30–35] have not included control groups, experimental controls, or randomized controlled studies. Only one study conducted by Wrzesien et al. (2011a) [29] has performed a randomized controlled study, comparing the in vivo exposure therapy with the AR exposure therapy.

3.5. AR and Acrophobia. In 2006, Juan et al. [43] advanced the use of immersive photography in an AR system to treat acrophobia. For evaluating this system, forty-one healthy volunteers walked around at the top of a staircase in a real environment and using the immersive photography environment. After their experience, the participants filled out the SUS questionnaire [49] to assess their subjective sense of presence. The data showed that the AR condition induced a sense of presence equal to the one experienced by the subjects in the real world.

Another study, conducted by Juan and Prez (2010) [36], compared an acrophobic virtual reality (VR) and AR environment assessing differences in the sense of presence and anxiety elicited by the two systems. Twenty healthy participants underwent both experimental conditions and after using each system (AR or VR), they completed an adapted SUS questionnaire [49]. Moreover, at six different moments during the two experiences, the participants were also asked to rate their anxiety level from 0 (no anxiety) to 10 (very high anxiety). Regarding the sense of presence, the results showed no differences between the two systems. Moreover, data revealed that anxiety levels decrease after the exposure.

3.6. Discussion. In the current scientific scenario, AR is a relevant issue and offers a viable alternative to traditional methods as the in vivo exposure therapy. This paper is the first systematic review that explores studies in the literature that use AR as a tool in the treatment of psychological disorders.

To date, the analysis of the literature suggested that AR has been mainly used in the evaluation and treatment of specific phobias such as phobias for small animals (cockroaches and spiders) and acrophobia. Regarding the assessment and treatment of specific phobias, from the literature analysis, it has been observed that thirteen studies have used an AR system in order to reduce anxiety, fear, and avoidance behaviours. Specifically, eleven studies concerned the evaluation and treatment of cockroaches and spiders phobia, while two studies affected the acrophobia.

As regards the cockroaches and spiders phobia, nine studies applied an AR-HMD system connected with a camera and a computer [28, 29, 31, 32, 34, 35, 37–40]. The camera was placed on the HMD helmet and connected with a USB to the computer where the AR system ran. The camera, recognizing the marker through the movement of the subject's head, projected the virtual images and objects in front of the subject. Among these studies, two [37, 38] compared different AR systems in order to evaluate their efficacy. In particular, Juan and Calatrava (2011) [37] compared an AR optical see-through (OST) system with a video see-through (VST), and Juan and Joele (2011) [38] contrasted an AR visible marker-based system with an AR invisible marker system.

The last studies have tested two innovative technological AR systems [30, 33]. In 2011, Botella et al. [33] have experimented with a new AR-mobile phone system and have created a mobile game, "Cockroach Game" [33] for the treatment of the cockroach phobia. The Cockroach Game is a puzzle game characterized by different levels of fear stimuli divided in two situations allowing players to progress in the game. The first included the possibility to see animals on various virtual situations inside the mobile phone. The second one allowed seeing them on real environment such as on the hands. In a further study, Wrzesien et al. [30] experiment with an innovative technological AR system, named therapeutic lamp (TL). The TL is an AR display projector created for the treatment of small animals' phobia.

As mentioned above, the majority of the studies concerned the assessment and the treatment of small animals phobias: five studies have focused on patients with cockroaches' phobia [32–35, 39, 40] and three on the cockroaches and spiders phobia [28, 29, 31] according the DSM-IV [27] criteria. Instead, three studies [30, 37, 38] have included healthy volunteers without any diagnosis of psychological and medical problems. In most of the studies [28, 29, 31–35], the AR exposure therapy was applied using the "one-session treatment" guidelines developed by Öst [41, 42].

A remarkable feature refers to the number of subjects included in these studies and the psychological and experiential measures in the assessment of the AR exposure therapy. Three studies experimented with one case study [33, 35, 39], two studies tested the AR system on six patients [32, 34], and one study assessed nine patients [31]. The remaining five studies have experimented with a higher number of subjects [28–30, 37, 38]. In 2011, Juan et al. [37, 38] assessed twenty-four healthy volunteers, while Wrzesien et al. (2011a; 2011b; 2013) [28–30], in the first two studies [28, 29], assessed twenty-two patients and, in the third study [30], they tested twenty-six healthy subjects.

Furthermore, all these studies considered anxiety as psychological measure showing that AR elicits anxiety as soon as the stimulus appeared that decreased during the time of exposure [28–35, 37–39]. Among these studies, seven of them also considered fear and avoidance behaviours showing that AR was able to reduce significantly fear and avoidance behaviours after the stimuli exposure [28–31, 33–35]. Finally, as regards the quality of user's experience, five studies measured presence [31, 32, 35, 37, 38] and three of them evaluated reality judgment [31, 32, 35] showing a high sense of presence in the AR system and reality judgment of the small animals.

Taking the two acrophobia studies [36, 43] into account, both applied an AR-HMD system to evaluate its efficacy comparing, in the first study, the AR-HMD system with a real environment [43] and, in the second study, the AR-HMD system with a VR- HMD system [36].

Regarding the number of subjects, both of the studies included healthy volunteers and, specifically, in 2006, Juan et al. [43] assessed forty-one participants, while in 2010 they evaluated twenty [36].

Among these studies, just Juan and Prez (2010) [36] considered anxiety as psychological measure showing that AR is effective as VR to induce anxiety when the stimulus appeared, which decreased during the exposure. In conclusion, both studies assessed the sense of presence showing, in the study of 2006 [43], a higher sense of presence in the AR environment than the real environment and in the study of 2010 [36] no difference between the AR and the VR environment.

The studies mentioned above suggest that AR offers many advantages such as the possibility to reproduce real objects in the environment (ecological validity), controlled situations, and ad hoc environments and objects in order to tailor the situations on the subject's needs and therapeutic purposes. However, the discussed studies are mostly preliminary researches that disclose some limitations due to the fact that AR exposure has only been recently tested for the evaluation and treatment of psychological disorders. The majority of the disclosed studies [28, 29, 31–35, 39] include a too small individual's sample and only one [29] of them is based on a randomized controlled design.

Instead, the reaming studies focused on testing the AR system on healthy subjects in order to verify the efficacy, usability, and the quality of user's experience of the new platform using randomized controlled studies [30, 36–38, 43].

Apart from these studies, the literature is lacking studies of AR on the treatments of other several psychological disorders.

Currently, other technological devices have proven effective in the evaluation and treatment of psychological disorders. In particular, VR has been demonstrated to be a very useful tool for the treatment of several psychological problems such as eating disorders and anxiety disorders [50–54]. The traditional treatment for these psychological disorders is the in vivo exposure therapy. As the term implies, the in vivo exposure therapy allows subjects to experience their fears under the close supervision of a physician or a therapist. However, not all patients benefit from this treatment and, according to Jefferey et al. (2000) [55] and Mann et al. (2007)

[56], some patients do not improve after treatment and others relapse in the long term.

In the last few years many studies have showed the efficacy of VR environments as therapeutic tool [50–53, 57–60]. A VR environment is a completely simulated three-dimensional environment modelled by a computer that allows, through the simulation, coping with critical and fear situation in a safe condition without losing sensory experience and physical presence.

A feature that makes VR a useful tool for evaluation and treatment of psychological disorders is to elicit emotional responses commensurate with the real ones.

The emotional engagement in virtual exposure depends on a number of factors such as the sense of presence. The sense of presence in virtual exposure is defined as the degree of "being there" in the virtual environment [61] and is marked by the sense of immersion and a sense of interaction. The sense of immersion is the result of the technological tools used such as the use of the HMD device that allows an immersive 3D experience to subjects. Instead, the sense of interaction is defined as the degree of interaction and manipulation for individuals of the virtual content or environment. Therefore, a high sense of presence in a virtual environment provides a greater realistic perception of the experience and consequently a strong and deep emotional engagement. This experience increases the ecological validity of the instrument of VR in the treatment of psychological disorders ensuring, through a simulated environment, a similar experience to the real one. At the same time, the user, feeling "present" in the simulated environment, perceives it as real and tends to transfer the expected skills from the virtual world to the real one in a nearly automatic manner.

Recent studies have shown that virtual exposure is as effective as in vivo exposure [62–64] and, in particular, the exposure to virtual environments has produced emotional and behavioural responses similar to those that occurred in the real world.

In the studies of Ferrer-García and Gutiérrez-Maldonado [62, 63] six virtual environments emotionally relevant and significant to subjects with eating disorders (ED) have been realized. Before and after exposure to each virtual environment, they estimated the state of anxiety and depression levels and the data showed that the virtual exposure is effective in causing and provoking relevant emotional responses to subjects.

Gorini et al. [64] started from the two previous studies but compared the virtual stimuli with the real ones and with their correspondent pictures to test the psychological (measuring the level of anxiety and sense of presence) and physiological (measuring the heart and respiration rate and the skin conductance) reactions to food in a sample of ED patients and healthy controls. After each experimental condition, in order to estimate the psychological variations, the subjects carried out two states of anxiety tests (STAI-S and VAS-A) and then the virtual exposure completed the presence questionnaire. The data showed that virtual stimuli are effective as the real one and more than photographs in eliciting emotional responses in ED patients and, more generally, the use of VR instead of real stimuli may simplify the framework of very specific contexts to help patients to cope with their conditions through a very controlled stimulation.

Finally, regarding the sense of presence, the results showed a significant degree of presence on the level of state anxiety in VR and real exposure conditions. As mentioned above, the VR gives the possibility of the subject to manipulate and interact with the environment and contents as in the real world.

Similarly, AR could be considered a useful tool in the evaluation and treatment of psychological disorders providing a number of advantages as the VR. Indeed, the AR, as VR simulation, can be seen as an experiential process, and the experience is an essential component in dealing with critical situations. Furthermore, this experience could allow exploring environments and situations hard to reproduce in reality and, as the VR, the feeling of presence in the AR system could permit the user to assign the learned behaviours from the AR world to the real one.

The AR simulation, therefore, could be considered an efficient method to act on real behaviours avoiding the risks and complications typical of real environments.

4. Conclusions

The aim of this paper was to review the recent studies on the use of AR in the evaluation and treatment of psychological disorders, focusing on current uses of AR in psychology and the various factors that make a new technique useful for the treatment of psychological disorders, expanding the possible fields of use of AR.

In general, the presented studies show that the AR seems to be a promising and useful tool for intervention in the treatment of specific phobias. Nevertheless, the small sample of subjects examined and the lack of control group and randomized controlled studies necessitate more randomized controlled experiments for exploring the AR efficacy in the clinical treatments. Despite these limitations, AR is proving to be a new technique useful to patients to experiment with technologically different and severe situations, as the exposure to fear or phobic stimuli, in a safe environment under the control of the therapist. Indeed, an AR system extends interactivity for assessing and supervising patient's reactions in real time and adaptability for creating controlled exposure settings based on the patient's needs or therapeutic purposes. Furthermore, it is to be noted that AR allows subjects/patients to manipulate and control the virtual elements, interacting with virtual objects placed in the real world in real time.

As a consequence, the experience to amplify the physical world with virtual contents can improve the ecological validity of the "mixed reality" [7] on environment, augmenting the sense of presence and engagement of the subject/patient. Indeed, studies of VR have shown that virtual stimuli are comparable to the real stimuli with regard to emotional responses [62–64]. Finally a strong and deep sense of presence and engagement can, also, improve the adherence to treatment.

Overall, AR may represent a new challenge for the assessment and treatment of different kinds of psychological disorders, such as eating and anxiety disorders performing new studies based on systematic measures of psychological and neurophysiological effects.

Abbreviations

AR:	Augmented Reality
BAT:	Behaviour avoidance test
FSQ:	Fear of Spiders Questionnaire
GPS:	Global Positioning System
HMD:	Head mounted display
OST:	Optical see-through
PRISMA:	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PRJQ:	Presence and Reality Judgment Questionnaire
SPBQ:	Spider Phobia Beliefs Questionnaire
SUDS:	Subjective Units of Discomfort Scale
TL:	Therapeutic lamp
VR:	Virtual reality
VST:	Video see-through.

Conflict of Interests

There is no financial or personal interest to report.

Authors' Contribution

Irene Alice Chicchi Giglioli, Federica Pallavicini, and Silvia Serino independently selected paper abstracts and titles and analyzed the full papers that met the inclusion criteria, resolving disagreements through consensus. Irene Alice Chicchi Giglioli, Federica Pallavicini, Elisa Pedroli, and Silvia Serino participated in the study design and wrote the paper. Elisa Pedroli, Silvia Serino, and Giuseppe Riva conceived the study and participated in its drafting, design, and coordination. All authors read and approved the final paper.

Acknowledgment

This work was supported by the Italian funded project "VRehab. Virtual Reality in the Assessment and TeleRehabilitation of Parkinson's Disease and Post-Stroke Disabilities" (RF-2009-1472190).

References

[1] O. Hugues, P. Fuchs, and O. Nannipieri, *New Augmented Reality Taxonomy: Technologies and Features of Augmented Environment*, Springer, New York, NY, USA, 2011.

[2] G. Riva, F. Davide, and W. A. IJsselsteijn, Eds., *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environment*, Studies in New Technologies and Practices in Communication, IOS Press, Amsterdam, The Netherlands, 2003.

[3] G. Riva, J. A. Waterworth, and E. L. Waterworth, "The layers of presence: a bio-cultural approach to understanding presence in natural and mediated environments," *Cyberpsychology and Behavior*, vol. 7, no. 4, pp. 402–416, 2004.

[4] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[5] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34–47, 2001.

[6] F. Zhou, H. B.-L. Dun, and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR," in *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pp. 193–202, IEEE, Cambridge, UK, September 2008.

[7] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Transactions on Information and Systems D*, vol. 77, pp. 1321–1329, 1994.

[8] P. Milgram, T. Haruo, U. Akira, and K. Fumio, "Augmented reality: a class of displays on the reality-virtuality continuum," in *Telemanipulator and Telepresence Technologies*, pp. 282–292, 1994.

[9] D. Yu, J. S. Jin, S. Luo, and W. Lai, "A useful visualization technique: a literature review for augmented reality and its application, limitation & future direction," in *Visual Information Communication*, M. L. Huang, Q. V. Nguyen, and K. Zhang, Eds., pp. 311–337, Springer, New York, NY, USA, 2010.

[10] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar, and N. Navab, "Marker-less tracking for AR: a learning-based approach," in *Proceedings of the International Symposium on Mixed And Augmented Reality (ISMAR '02)*, pp. 295–304, IEEE Computer Society, Darmstadt, Germany, 2002.

[11] S. de Buck, F. Maes, J. Ector et al., "An augmented reality system for patient-specific guidance of cardiac catheter ablation procedures," *IEEE Transactions on Medical Imaging*, vol. 24, no. 11, pp. 1512–1524, 2005.

[12] C. S. Özbek, B. Giesler, and R. Dillmann, "Jedi training: playful evaluation of head-mounted augmented reality display systems," in *The Conference Medical Imaging*, Proceedings of SPIE, San Diego, Calif, USA, May 2004.

[13] B. Schwald and B. Laval, "An augmented reality system for training and assistance to maintenance in the industrial context," in *Proceedings of the International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, pp. 425–432, University of West Bohemia, Plzen, Czech Republic, 2003.

[14] R. Grasset, X. Decoret, and J. D. Gascuel, "Augmented reality collaborative environment: calibration and interactive science editing," in *Proceedings of the Virtual Reality International Conference (VRIC '01)*, Laval Virtual, 2001.

[15] T. N. Arvanitis, A. Petrou, J. F. Knight et al., "Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities," *Personal and Ubiquitous Computing*, vol. 13, no. 3, pp. 243–250, 2009.

[16] L. Kerawalla, R. Luckin, S. Seljeflot, and A. Woolard, "'Making it real': exploring the potential of augmented reality for teaching primary school science," *Virtual Reality*, vol. 10, no. 3-4, pp. 163–174, 2006.

[17] Y.-J. Chang, Y.-S. Kang, and P.-C. Huang, "An augmented reality (AR)-based vocational task prompting system for people with cognitive impairments," *Research in Developmental Disabilities*, vol. 34, no. 10, pp. 3049–3056, 2013.

[18] R. Hervás, J. Bravo, and J. Fontecha, "An assistive navigation system based on augmented reality and context awareness for people with mild cognitive impairments," *IEEE Journal of*

Biomedical and Health Informatics, vol. 18, no. 1, pp. 368–374, 2014.

[19] G. A. Assis, A. G. Corrêa, M. B. R. Martins, W. G. Pedrozo, and R. D. Lopes, "An augmented reality system for upper-limb post-stroke motor rehabilitation: a feasibility study," *Disability and Rehabilitation: Assistive Technology*, vol. 4, pp. 1–8, 2014.

[20] A. J. Espay, Y. Baram, A. K. Dwivedi et al., "At-home training with closed-loop augmented-reality cueing device for improving gait in patients with Parkinson disease," *Journal of Rehabilitation Research and Development*, vol. 47, no. 6, pp. 573–582, 2010.

[21] X. Luo, T. Kline, H. C. Fischer, K. A. Stubblefield, R. V. Kenyon, and D. G. Kamper, "Integration of augmented reality and assistive devices for post-stroke hand opening rehabilitation," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS '05)*, pp. 6855–6858, IEEE, September 2005.

[22] M. Markovic, S. Dosen, C. Cipriani, D. Popovic, and D. Farina, "Stereovision and augmented reality for closed-loop control of grasping in hand prostheses," *Journal of Neural Engineering*, vol. 11, no. 4, Article ID 046001, 2014.

[23] M. Khademi, H. M. Hondori, L. Dodakian, S. Cramer, and C. V. Lopes, "Comparing 'pick and place' task in spatial augmented reality versus non-immersive virtual reality for rehabilitation setting," in *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '13)*, pp. 4613–4616, 2013.

[24] H. M. Hondori, M. Khademi, L. Dodakian, S. C. Cramer, and C. V. Lopes, "A Spatial Augmented Reality rehab system for post-stroke hand rehabilitation," *Studies in Health Technology and Informatics*, vol. 184, pp. 279–285, 2013.

[25] O. Baus and S. Bouchard, "Moving from virtual reality exposure-based therapy to augmented reality exposure-based therapy: a review," *Frontiers in Human Neuroscience*, vol. 8, article 112, 2014.

[26] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and P. Grp, "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," *PLoS Medicine*, vol. 6, no. 7, Article ID e1000097, 2009.

[27] A. P. Association, *Diagnostic and Statistical Manual for Mental Disorders (DSM-IV-TR)*, American Psychiatric Association, American Occupational Therapy, Washington, DC, USA, 4th edition, 2000.

[28] M. Wrzesien, J.-M. Burkhardt, M. A. Raya, and C. Botella, "Mixing psychology and HCI in evaluation of augmented reality mental health technology," in *Proceedings of the 29th Annual CHI Conference on Human Factors in Computing Systems (CHI '11)*, pp. 2119–2124, Vancouver, Canada, May 2011.

[29] M. Wrzesien, J.-M. Burkhardt, M. Alcan~iz, and C. Botella, "How technology influences the therapeutic process: a comparative field evaluation of augmented reality and in vivo exposure therapy for phobia of small animals," in *Proceedings of the Human-Computer Interaction (INTERACT '11)*, pp. 523–540, Lisbon, Portugal, 2011.

[30] M. Wrzesien, M. Alcañiz, C. Botella et al., "The therapeutic lamp: treating small-animal phobias," *IEEE Computer Graphics and Applications*, vol. 33, no. 1, pp. 80–86, 2013.

[31] M. C. Juan, M. Alcañiz, C. Monserrat, C. Botella, R. M. Baños, and B. Guerrero, "Using augmented reality to treat phobias," *IEEE Computer Graphics and Applications*, vol. 25, no. 6, pp. 31–37, 2005.

[32] J. Breton-López, S. Quero, C. Botella, A. García-Palacios, R. M. Baños, and M. Alcañiz, "An augmented reality system validation for the treatment of cockroach phobia," *Cyberpsychology, Behavior, and Social Networking*, vol. 13, pp. 705–710, 2010.

[33] C. Botella, J. Breton-López, S. Quero et al., "Treating cockroach phobia using a serious game on a mobile phone and augmented reality exposure: a single case study," *Computers in Human Behavior*, vol. 27, no. 1, pp. 217–227, 2011.

[34] C. Botella, J. Bretón-López, S. Quero, R. Baños, and A. García-Palacios, "Treating cockroach phobia with augmented reality," *Behavior Therapy*, vol. 41, no. 3, pp. 401–413, 2010.

[35] C. Botella, R. Banos, B. Guerrero, M. Juan, and M. Alcañiz, "Mixing realities? An augmented reality system for the treatment of cockroach phobia," *CyberPsychology & Behavior*, vol. 8, pp. 305–306, 2005.

[36] M. C. Juan and D. Prez, "Using augmented and virtual reality for the development of acrophobic scenarios. Comparison of the levels of presence and anxiety," *Computers and Graphics*, vol. 34, no. 6, pp. 756–766, 2010.

[37] M. C. Juan and J. Calatrava, "An augmented reality system for the treatment of phobia to small animals viewed via an optical see-through HMD. Comparison with a similar system viewed via a video see-through HMD," *International Journal of Human-Computer Interaction*, vol. 27, no. 5, pp. 436–449, 2011.

[38] M. C. Juan and D. Joele, "A comparative study of the sense of presence and anxiety in an invisible marker versus a marker augmented reality system for the treatment of phobia towards small animals," *International Journal of Human Computer Studies*, vol. 69, no. 6, pp. 440–453, 2011.

[39] M. C. Juan, C. Botella, M. Alcañiz et al., "An augmented reality system for treating psychological disorders: application to phobia to cockroaches," in *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '04)*, pp. 256–257, IEEE, November 2004.

[40] C. M. Botella, M. C. Juan, R. M. Baños, M. Alcañiz, V. Guillén, and B. Rey, "Mixing realities? An application of augmented reality for the treatment of cockroach phobia," *Cyberpsychology & Behavior*, vol. 8, no. 2, pp. 162–171, 2005.

[41] L.-G. Ost, P. M. Salkovskis, and K. Hellstrom, "One-session therapist-directed exposure vs. self-exposure in the treatment of spider phobia," *Behavior Therapy*, vol. 22, no. 3, pp. 407–422, 1991.

[42] L. G. Öst, *Rapid Treatment of Specific Phobias*, John Wiley & Sons, New York, NY, USA, 2000.

[43] M. C. Juan, R. Baños, C. Botella, D. Pérez, M. Alcañiz, and C. Monserrat, "An augmented reality system for the treatment of acrophobia: the sense of presence using immersive photography," *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 4, pp. 393–402, 2006.

[44] J. Wolpe, *The Practice of Behavior Therapy*, Pergamon Press, New York, NY, USA, 1969.

[45] I. M. Marks and A. M. Mathews, "Case histories and shorter communication," *Behaviour Research and Therapy*, vol. 17, pp. 263–267, 1979.

[46] J. Szymanski and W. J. O'Donohue, "Fear of spiders questionnaire," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 26, no. 1, pp. 31–34, 1995.

[47] A. Arntz, E. Lavy, G. van Den Berg, and S. van Rijsoort, "Negative beliefs of spider phobics: a psychometric evaluation of the spider phobia beliefs questionnaire," *Advances in Behaviour Research and Therapy*, vol. 15, no. 4, pp. 257–277, 1993.

[48] R. M. Ban~os, S. Quero, S. Salvador, and C. Botella, *The Role of Presence and Reality Judgement in Virtual Environments in Clinical Psychology*, Verlag Integrative Psychiatrie, Innsbruck, Austria, 2005.

[49] M. Slater, M. Usoh, and A. Steed, "Depth of presence in virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 3, pp. 130–144, 1994.

[50] A. Gorini and G. Riva, "The potential of Virtual Reality as anxiety management tool: a randomized controlled study in a sample of patients affected by generalized anxiety disorder," *Trials*, vol. 9, article 25, 2008.

[51] T. D. Parsons and A. A. Rizzo, "Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 39, no. 3, pp. 250–261, 2008.

[52] G. Riva, "Virtual reality in eating disorders and obesity: state of the art and future directions," *CyberPsychology & Behavior*, vol. 8, p. 351, 2005.

[53] G. Riva, M. Bacchetta, M. Baruffi, and E. Molinari, "Virtual reality-based multidimensional therapy for the treatment of body image disturbances in obesity: a controlled study," *Cyberpsychology and Behavior*, vol. 4, no. 4, pp. 511–526, 2001.

[54] A. Rizzo, J. C. Buckwalter, and C. Van der Zaag, "Virtual environment applications in clinical neuropsychology," in *The Handbook of Virtual Environments*, K. M. Stanney, Ed., pp. 1027–1064, Erlbaum, New York, NY, USA, 2002.

[55] R. W. Jeffery, L. H. Epstein, G. T. Wilson et al., "Long-term maintenance of weight loss: current status," *Health Psychology*, vol. 19, no. 1, pp. 5–16, 2000.

[56] T. Mann, A. J. Tomiyama, E. Westling, A.-M. Lew, B. Samuels, and J. Chatman, "Medicare's search for effective obesity treatments. Diets are not the answer," *American Psychologist*, vol. 62, no. 3, pp. 220–233, 2007.

[57] G. M. Manzoni, G. L. Cesa, D. Villani, G. Castelnuovo, E. Molinari, and G. Riva, "VR-enhanced treatment of anxiety in obese subjects: a follow-up study on trait-anxiety, psychological symptornatology, and generalized self-efficacy," *CyberPsychology & Behavior*, vol. 9, pp. 699–700, 2006.

[58] S. Côté and S. Bouchard, "Virtual reality exposure for phobias: a critical review," *Journal of Cyber Therapy & Rehabilitation*, vol. 1, no. 1, pp. 75–91, 2008.

[59] F. Pallavicini, A. Gaggioli, S. Raspelli et al., "Interreality for the management and training of psychological stress: study protocol for a randomized controlled trial," *Trials*, vol. 14, no. 1, article 191, 2013.

[60] A. Gaggioli, G. Pioggia, G. Tartarisco et al., "A system for automatic detection of momentary stress in naturalistic settings," *Studies in Health Technology and Informatics*, vol. 181, pp. 182–186, 2012.

[61] W. Barfield, D. Zeltzer, T. B. Sheridan, and M. Slater, *Virtual Environments and Advanced Interface Design*, Oxford University Press, Oxford, UK, 1995.

[62] M. Ferrer-García, J. Gutiérrez-Maldonado, A. Caqueo-Urízar, and E. Moreno, "The validity of virtual environments for eliciting emotional responses in patients with eating disorders and in controls," *Behavior Modification*, vol. 33, no. 6, pp. 830–854, 2009.

[63] J. Gutiérrez-Maldonado, M. Ferrer-García, A. Caqueo-Urízar, and A. Letosa-Porta, "Assessment of emotional reactivity produced by exposure to virtual environments in patients with eating disorders," *Cyberpsychology and Behavior*, vol. 9, no. 5, pp. 507–513, 2006.

[64] A. Gorini, E. Griez, A. Petrova, and G. Riva, "Assessment of the emotional responses produced by exposure to real food, virtual food and photographs of food in patients affected by eating disorders," *Annals of General Psychiatry*, vol. 9, article 30, 2010.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 358638, 15 pages
http://dx.doi.org/10.1155/2015/358638

Research Article

Novel Virtual User Models of Mild Cognitive Impairment for Simulating Dementia

Sofia Segkouli,[1,2] Ioannis Paliokas,[1] Dimitrios Tzovaras,[1] Thanos Tsakiris,[1] Magda Tsolaki,[3] and Charalampos Karagiannidis[2]

[1] *Centre for Research and Technology Hellas (CERTH), Information Technologies Institute (ITI), P.O. Box 60361, 6th Charilaou-Thermi, 57001 Thessaloniki, Greece*
[2] *Department of Special Education, University of Thessaly, Argonafton & Filellinon Street, 38221 Volos, Greece*
[3] *Pan-Hellenic Federation of Alzheimer's disease and Related Disorders, 13 P. Syndika Street, 546 43 Thessaloniki, Greece*

Correspondence should be addressed to Sofia Segkouli; sofia@iti.gr

Received 10 October 2014; Revised 23 January 2015; Accepted 4 February 2015

Academic Editor: Pietro Cipresso

Virtual user modeling research has attempted to address critical issues of human-computer interaction (HCI) such as usability and utility through a large number of analytic, usability-oriented approaches as cognitive models in order to provide users with experiences fitting to their specific needs. However, there is demand for more specific modules embodied in cognitive architecture that will detect abnormal cognitive decline across new synthetic task environments. Also, accessibility evaluation of graphical user interfaces (GUIs) requires considerable effort for enhancing ICT products accessibility for older adults. The main aim of this study is to develop and test virtual user models (VUM) simulating mild cognitive impairment (MCI) through novel specific modules, embodied at cognitive models and defined by estimations of cognitive parameters. Well-established MCI detection tests assessed users' cognition, elaborated their ability to perform multitasks, and monitored the performance of infotainment related tasks to provide more accurate simulation results on existing conceptual frameworks and enhanced predictive validity in interfaces' design supported by increased tasks' complexity to capture a more detailed profile of users' capabilities and limitations. The final outcome is a more robust cognitive prediction model, accurately fitted to human data to be used for more reliable interfaces' evaluation through simulation on the basis of virtual models of MCI users.

1. Introduction

As people age, they experience alterations in cognitive abilities such as working memory, at attention processes and spatial cognition, especially when confronting new technology [1]. The design of human-computer interfaces' requires careful study of multitasking capability [2] and has to take into account cognitive aging changes on older adults' perception [3]. The line between normal aging and dementia may comprise conditions in which heterogeneous patterns of cognitive impairment may be observed. Indeed, memory disorders with no dementia in the elderly population are frequently reported, and their prevalence varies from 22% to 56% [4].

So far, research efforts have developed cognitive architectures and theories in order to capture the essential representations of cognition. This activity has gradually moved from a focus on the functional capabilities of architectures to the ability to model the details of human behaviour, and, more recently, brain activity [5]. Cognitive models have been used to simulate humans' performing multiple tasks in order to improve the quality and usability of the interface designs [6]. A bunch of models that merge psychology and artificial intelligence termed as cognitive architecture like SOAR [7], ACT [8], and EPIC [9] have been used to simulate human-machine interaction to both explain and predict interaction behaviour. A simplified view of these cognitive architectures is known as the GOMS model [10].

Commonly, models developed using cognitive architectures consider the uncertainty of human behavior in detail but have not been widely adopted for simulating HCI as their use demands a detailed knowledge of psychology [11]. Existing user modeling tools, focusing on inclusive interaction like EASE [12] or CogTool [13], do not yet cover a wide range of users with perceptual, cognitive, and motor disabilities.

The SOAR architecture was created to demonstrate general human-level intelligence and focuses more on high-level functionality than on low-level cognitive fidelity, which makes it less suited to predict people's errors and limitations [14]. EPIC is especially suited for modeling human multimodal and multiple-task performance including sensory-motor processors. The ACT-R theory for simulating and understanding human cognition enabled serious consideration of the cognitive, perceptual, and motor capabilities of the user, to describe interactivity in HCI [15] and provided a reliable cognitive architecture for the development of large-scale, functional, cognitively motivated models. ACT-R implementations may vary in technical details such as time parameters of the user model [16]. For computational reasons, most actual ACT-R implementations use additional submodules as noise generators applied on the output in order to create noisy parameters.

The cognitive module proposed in this paper is based on a modification of the ACT-R implementation model in order to achieve more accurate prediction of the MCI elderly memory changes mirrored at the performance on infotainment tasks. One effective approach could be to intervene in the internal part of the ACT-R cognitive architecture by introducing a dual submodule (one for the declarative memory and one for the procedural) that can affect memory as well as the overall cognitive functionality. The further challenge was to determine the error rates and the extra time needed by MCI virtual user models (VUMs) and compare timing and success rate results with those generated by real MCI users while interacting with interfaces.

(1.1) Cognitive VUMs Based on ACT-R. Cognitive models have been used to improve the quality of interface design by applying what is known from psychology to the design of interfaces [6]. ACT-R, EPIC, and SOAR have been employed extensively in psychology and the cognitive sciences to model human behaviour [17]. Moreover, these techniques may be used to explain and predict human-computer interaction. However these cognitive architectures, known as the GOMS (goals, operators, methods, and selection rules) model, are mainly suitable for modeling the optimal (skilled) behaviour of users.

Various user model representations have been proposed from ontology based architectures [18] to XML-based languages [19]. Based on ACT-R cognitive architecture, Serna et al. [20] aimed to model and simulate cognitively impaired people such those suffering from Alzheimer's disease by evaluating performance in the execution of an activity of daily living. They simulated the loss of memory and increase in error for a representative task at kitchen by changing different ACT-R parameters. The technique is interesting but their model still needs rigorous validation through other tasks and user communities.

Biswas et al. [21] developed a simulator to help with the evaluation of assistive interfaces using the CPM_GOMS [10] model to simulate the optimal behavior and a new model based on Markov processes for suboptimal behavior. The CogTool system [22] builds on the existing architecture of GOMS models and ACT-R system in order to provide quantitative prediction on interaction. However, the system simulates expert performance with GOMS modeling support while the ACT-R system helps to simulate novice users. It does not yet seem to be used for users with disability or assistive interaction techniques. The GUIDE user modeling system considers the needs of users with disability and age related impairments. Despite the many conceptual advances in cognitive architecture research and the strength in modeling cognition and general intelligence, there are some open issues that deserve attention. Among their known limitations is that detailed physical, cognitive, and behavioral characteristics cannot be sufficiently represented.

Comparatively little research has been paid on the simulation of memory loss and more rigorous approaches of predictive parameters that concern specific population groups such as the mild cognitive impairment are needed. Virtual user models have commonly been shown to provide insufficient knowledge to the product designers in order to differentiate normal from pathological origins of age-related cognitive changes and provide interface prototypes with enhanced accessibility.

(1.2) Novel MCI Virtual User Models. In our approach, ACT-R has been chosen as the theoretical framework as it is the most representative example of cognitive modeling architecture being pushed to its limits by large-scale modeling and used to develop specific models. ACT-R comprises a series of modules which serve to represent different information such as goals and declarative memory. ACT-R can model characteristic errors of human performance; however, the work in cognitive simulation is in progress and in continuous update to reflect subtle changes of cognition and behavior. The specific challenge was to extend capabilities and functionality of already existing cognitive mechanisms. Therefore the proposed dual submodule could be viewed as a shift in specific cognitive representations. Moreover different cognitive modules could be plugged in for more accurate evaluation purposes.

Initially, our main challenge was to enhance the cognitive realism of computer generated models and afterwards provide a second generation of VUMs on the basis of a novel submodule that will reflect representative and subtle cognitive changes of elderly and people with memory decline as an early onset in mild cognitive impairments. MCI is the transitional stage between normal ageing and dementia. More specifically, amnestic mild cognitive impairment (aMCI) is characterised by memory complaints, so the simulated memory of the VUM had to be affected. Moreover, verbal fluency ability is affected in a MCI and thus some extra delays in communication tasks would give VUM results closer to those of the actual users [23].

In order to describe a set of users with specific disabilities we instantiated a generic VUM (GVUM), representing the specific group of elderly and MCI population. The general concept of the generic user model has been defined by the Veritas project VUMs [24]. In this experiment, as in previous studies [25], these instances of the GVUMs were intended to be used for simulation purposes to test user interfaces against various disabilities. In order to form and evaluate the proposed virtual user models we have been based on high level languages for describing cognitive models and simulation frameworks supporting these languages. More specifically, our virtual user models were formed by separating user models from task models according to the UsiXML Language [26].

Our simulator, based on well-known literature findings [25, 27], like the Hick-Hyman law for cognitive information capacity [28, 29] and Fitts' law for human movement [30] became capable of predicting the number of interaction events within infotainment scenarios and the time needed to complete them [31]. Hick's law states that the time between a stimulus and the user's response is increasing by the number of available choices at any given time. Users, when facing a decision making problem, need more time to react as their perceived alternatives increase. In a GUI control element, for instance, a long drop-down list requires more time to pick the wished choice. The time is calculated in the beginning of the user's reaction in order not to be confused with motor abilities. Hick's law was not used in equally probable choices; the following formula was used to take into account the entropy of the decision:

$$H = \sum_{i}^{n} p_i \log_2 \left(\frac{1}{p_i} + 1 \right), \tag{1}$$

where p_i is the probability of the ith alternative (information theoretic entropy). Fitt's law is a well-established model which describes the time needed for a human movement from an initial state to a target based on the (a) Euclidean distance to the target and (b) the size of the target. In simple words, this model is used to predict pointing actions either by a physical touch or a virtual one, like in cursor pointing using a mouse device. The predicted time is given by the following equation:

$$T = a + b \log_2 \frac{2D}{W}, \tag{2}$$

where: T: time needed to complete the task, a, b: model parameters, D: distance from the current cursor position to the center of the target, and W: width of the target along the axis of motion.

To overcome the limitation of screen sizes, the size of the target W is considered very large (infinite) in case of targets located at the edge of the screen. Task analysis and prediction of real user's performance (durations) have been used in the past with good prediction accuracy over predefined tasks [32].

This study investigated how subtle changes in cognition could affect GUI interaction for elderly and MCI groups. Cognitive functioning data correlating with memory capacity have been considered to be of critical importance to be encoded into VUMs that would be assessed through simulations over GUIs. Figure 1 briefly describes the submodules of the new ACT-R approach. The first line contains memory components which are being affected by cognitive impairments. The control component located at the center (procedural) orchestrates all others by exchanging messages. Located at the bottom are the perceptual system possibly affected by visual and/or aural impairments and the response system where the motion planning of the cursor is generated, along with the button clicks and the key strokes (motor). The extension to the existing architecture is the MCI modeling submodule which temporarily modifies the random rules, memory retrievals, and messages exchanged between other parts of the VUM, to simulate the cognitive decline. An internal interface has been designed for this submodule. For compatibility reasons the functionality of the submodules has been implemented in the simulator with keeping the VUMs structures unaffected.

The modification of the memory-related features of the VUM (mostly declarative memory and current goal) is simulated by inserting additional time delays on the performance of the elderly VUM. Thus, according to the current approach, the time to complete a given scenario by the MCI cognitively impaired VUMs is defined by

(i) the human movement model known as Fitt's law mounted in the motor sector of the response system,

(ii) Hick's law which describes the time needed by the VUM to make a choice among many interaction controls on a given interface based on its current goal [28],

(iii) the new additional time delay inserted in the memory modules caused by the MCI and being scaled according to the assumed Montreal Cognitive Assessment test (MoCa) [33] and the Boston Naming Test (BNT) [34] results,

(iv) other delays issues to the ACT-R itself, the hypothetical age of the VUM, or the existence of physical impairments.

Summarizing, the proposed MCI submodule is affecting the error rate and the execution time of tasks. Those two parameters were trained using tests with real users that are reported in the next section.

2. General Objectives

Computer and information technologies' use could improve the quality of life of older people. However, successful use of technology by the elderly is limited by age-related cognitive and perceptual changes. Designers have to use as an index not generally age but concrete cognitive, perceptual, and behavioural changes to identify early indicators of dementia and customize the entire process of design accordingly [1]. Existing design practices have provided a few accessibility features for elderly or disabled users, mainly covering motor, vision, and hearing impairments. Unfortunately, up to now accessibility guidelines do not systematically consider users with perceptual and cognitive abilities [21].

Within this research work, we undertook a multifacet study, focusing on robust cognitive prediction models for

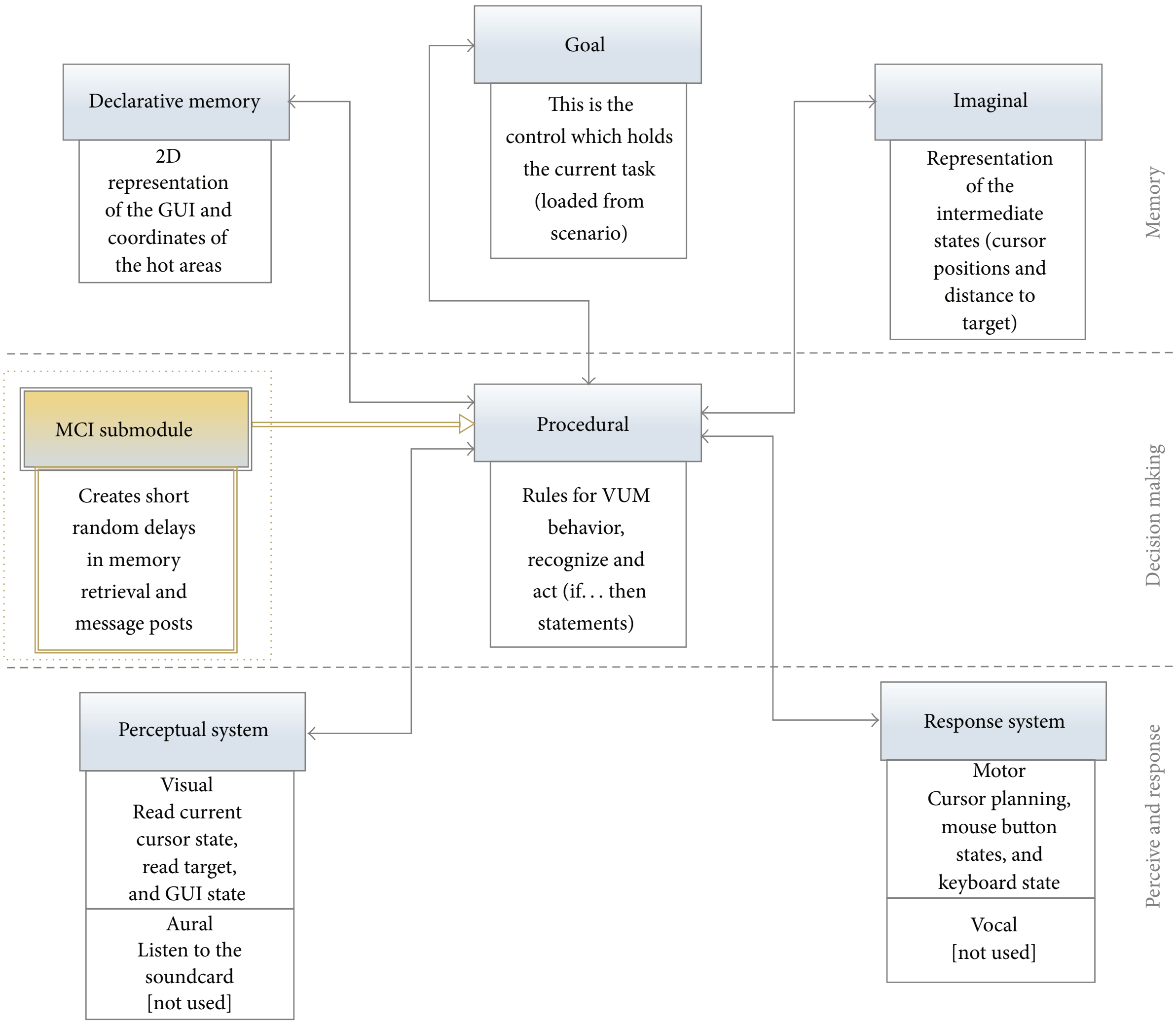


Figure 1: The revised cognitive model of the VUMs (extended ACT-R model).

mild cognitive impairment within the scope of graphical users' interfaces accessibility and evaluation. The main aim is to introduce novel virtual user models with enhanced predictive validity in mental processes that will be utilised for accurate simulation results in interface design.

Cognitive-based evaluation of interface accessibility and evaluation are proposed as interface design support towards more inclusive graphical user interfaces. However, user interface evaluation is also the means to arrive at the optimization of the proposed novel cognitive models. Our approach is to initially research the relationship between some design variables and the mental processes required to perform successfully some indicated scenarios and then utilise the effectiveness of user interfaces to optimize the MCI VUMs. The main vision of the proposed research is approached through the synthesis of the following objectives.

(i) Model elderly with mild cognitive impairment for an in-depth analysis of their perceptual and cognitive abilities.

(ii) Evaluate cognitive decline through GUI accessibility assessment.

(iii) Utilize the effectiveness of user interfaces to optimize MCI VUMs.

To that end, we organized our work in a set of 4 studies, each one aiming to provide a specific set of qualitative and quantitative results in order to arrive at the aforementioned objectives. These studies, as well as their methodologies and results, are detailed in the following sections.

3. Study 1: Preliminary Study with Real Subjects

3.1. Introduction to Study 1. With an ever-increasing proportion of the total elderly population suffering from MCI symptoms, GUI designers need to deliver ICT products which have been previously tested for their accessibility features. According to our approach, simulation testing can make this possible to a certain extent using VUMs instead of real users. To train the MCI cognitive VUMs, that is, to stabilize the cognitive-impaired VUM parameters and regulate their simulating performance, we first needed to assess the MCI symptoms of real users and then examine these users' performance (extra time needed to complete tasks and a higher error rate). To achieve this, initially real users have been recruited to perform well-established screening tests for cognitive evaluation.

3.2. Materials and Methods of Study 1. Among various MCI screening tests that are already applied in day-care organizations, the short version of BNT screening test (30BNT) was chosen as one of the most well-established MCI tests. The MoCa test, supplementary to BNT, was also chosen due to its verbal fluency part so as to test users' verbal ability and strengthen their profiles. A computer-based version of the aforementioned tests was implemented and its validity has been checked against the paper-and-pencil versions of the tests.

The test had five discrete steps: (a) introduction to the scopes of the test by the pilot test supervisor (5 min), (b) initialization of hardware and software (3 min), (c) creating new user profile (1 min), (d) performing a demo example for each test (30 sec), and (e) actual test and log file creation (20 min). During the screening tests, the cognitively impaired elderly, as well as all other elderly and young participants, did not have to navigate themselves in the screening application. Instead, a member of the research personnel (RP) was navigating for them; from the participant's log files a constant time delay factor (psychomotor noise) was removed to correct the timestamps. A simple reaction time (SRT) was taken as a constant calculated from Hick-Hyman law [28, 29].

Considering the right and wrong answers of the elderly as equally probable choices, the RP had to press the "correct answer" button or press the spacebar to start writing the user's own word. Each RP, who was responsible for pilot tests, performed ten ($N = 10$) tries on a reaction time test to measure the response times for keyboard presses and mouse clicks. The average response times were removed from the time-stamped answers of all elderly testers who were recorded by this RP. Regarding the BNT test, the subjects who were unable to provide the correct name of the depicted object within 20 s were given a semantic cue and if they were still unable to give the answer after a further 20 sec, they were given a verbal cue. The 20 sec threshold for giving help was given by the institution responsible for taking the paper-based BNT test.

Finally, the screening assessment produced log files which contained the time-stamped responses of the participants. A mass log reader tool was used later for data extraction. The screening test itself, as well as the log files reader and other supplementary software tools, were engineered with Delphi and C++ programming languages. A screenshot of the computer-based MCI screening test interface can be seen in Figure 2.

MCI screening tests were performed with real users, constituted by three groups: (a) young people, (b) healthy elderly, and (c) MCI users (aMCI). Recruitment criteria for

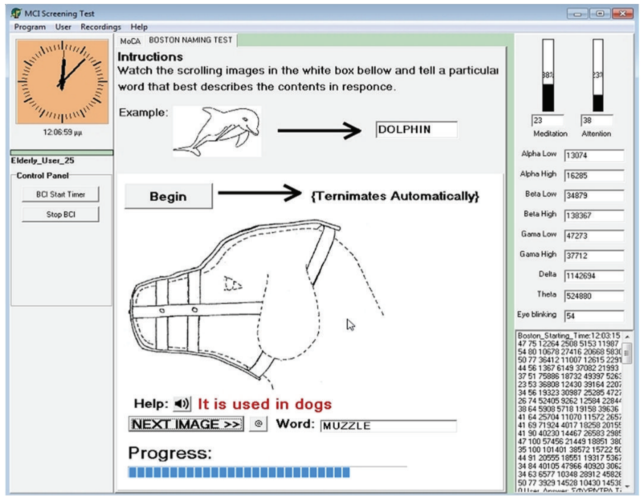


FIGURE 2: The computer-based BNT Interface.

MCI users included at least 55 years of age, being self-handlers, having basic computer skills and fluency in Greek language (checked by phone interview). Twenty-five ($N = 25$) elderly persons of different ages from 55 to 78 ($M = 65.56$, $SD = 5.89$) composed the testing group. Approximately 40% ($N = 10$) of the elderly were MCI positives, nondemented. The control group consisted of eleven ($N = 11$) young expert computer users, 26.27 years old in average ($SD = 1.95$).

3.3. Results of Study 1. The data produced by the computer-based version of the MCI screening tests showed significantly strong correlation with those produced by the institutionalized hard-copy versions of the 30BNT and MoCa. This initial step was necessary to make the computer-based versions of the MCI screening tests a valid reference for the following steps. To be noted that the follow-up time of BNT test was six (6) months later in average for most participants and less than one (1) year for all people taking part in this study.

The linguistic MoCa results were found normally distributed (Shapiro-Wilk test), in contrast with BNT results which were not, since results were significantly negatively skewed. Linguistic MoCa results for healthy elderly ($M = 13.80$, $SD = 4.94$) were found significantly different ($t = 2.317$, $P < 0.05$) than those having MCI ($M = 9.10$, $SD = 4.84$). The Shapiro-Wilk test showed significant deviance from normality for BNT results, so nonparametric test was used. The Mann-Whitney U test gave $P < 0.05$ for the elderly group. So, among elderly users, those having MCI ($M = 21.78$, $SD = 5.40$) had significantly different BNT results than the healthy ones ($M = 25.90$, $SD = 1.97$) as expected. Regarding the age variable, the independent samples test showed that, as expected, no significant differences were found on linguistic MoCa results between the elderly users and the young group ($P > 0.05$) [35]. Similar results were found for BNT scores: a Kruskal-Wallis test confirmed that age does not make the difference ($P > 0.05$).

The Mann-Whitney U test for the duration of the computer-based BNT test, a variable that was found not

Table 1: MCI screening test results.

Variables	Young	Healthy elderly	MCI
Sex M/F	4/7	10/5	6/4
Age (in years)	26.27 (SD = 1.95)	64.69 (SD = 4.80)	66.70 (SD = 7.18)
Education (in years)	17.36 (SD = 0.80)	13.46 (SD = 3.33)	12.9 (SD = 3.38)
MoCa number of words (in 60 sec)	13.64 (SD = 4.03)	13.80 (SD = 4.94)	9.1 (SD = 4.84)
MoCa duration between words (in sec)	4.70 (SD = 1.186)	5.05 (SD = 2.36)	8.45 (SD = 4.76)
BNT answers without help	25.56 (SD = 2.18)	25.90 (SD = 1.97)	21.78 (SD = 5.40)
BNT answers with semantic help	26.00 (SD = 2.06)	26.20 (SD = 1.39)	22.22 (SD = 5.12)
BNT answers with phonemic help	26.11 (SD = 2.14)	26.20 (SD = 1.39)	22.56 (SD = 5)
BNT duration (in sec)	113.3 (SD = 33.60)	117.49 (SD = 21.80)	167.3 (SD = 71.65)

normally distributed, showed that the time needed by participants to complete the BNT is not an efficient separator for MCI users ($P > 0.05$) and only the right-responses are. Table 1 presents the results per young, well-being elderly and MCI groups. Research results confirm that among young and healthy elderly the decline in lexical retrieval is not notable [35]. Finally, variables of the linguistic MoCa and the BNT present a notable correlation ($r = 0.395$, $P \leq 0.05$), though further analysis revealed different qualities among them. Furthermore, statistical analysis reported good cohesion and separation ability of the linguistic MoCa and BNT clustering model with a score of about ≈ 0.6.

3.4. Brief Discussion of Study 1 Results. The screening procedure through MoCa and BNT tests serves as a starting point to the whole methodology of the current research work. Initially, the MCI symptoms of real users had to be assessed as evidence of cognitive decline. This evidence was needed to formulate the quantitative data which will be later related to users' performance while performing interface tasks. Thus, results from MoCa and BNT screening tests assessed users' cognitive and perceptual factors that encompass also the mental state behind response time and general behavioral performance that will be resulted from the following study.

4. Study 2: Simulating Tasks with Cognitive VUMs

4.1. Introduction to Study 2. Our proposed VUMs were trained based on the data obtained from the afore-described screening tests. The trained MCI cognitive VUMs were then used for the simulation of tasks execution in human-computer interaction with graphical user interfaces (GUI). VUMs in this paper are named according to the cognitive impairment they simulate and the strength (percentile of the total population) of the impairment they correspond to. For example, the "75MCI" VUM is a virtual user model which can simulate the behavior of the 75% of the cognitive impaired elderly population, which is calculated by creating a Gaussian distribution from the obtained data and taking the 75th percentile (i.e., the distribution's value, below which 75% of the observations may be found). The alphanumeric suffix describes the impairment and the numeric prefix indicates the strength or the earliest symptomatic stage of MCI.

To obtain a reference for comparing the performance of the virtual users (i.e., trained VUMs) to the performance of real users, initially, the testing GUI application is initiated in the host device (PC, laptop, smartphone, PDA, and tablet) and the optimal user (OpUs), who is a nonimpaired real young user with expert computer handling abilities and sufficient knowledge of the testing interface, is asked to perform a series of tasks. Those groups of tasks, called scenarios, are described by the task model specification (an XSD schema) and are stored in the task model bank. The optimal recording is critical as it provides a proper reference (ground truth) for comparing the performance of VUMs to the one of real users at later stages. Before any simulation preparation, GUI designers need to define the activities they want to test for accessibility. This means that a formal description of the series of actions in order to arrive at a given result must be produced. To proceed to the actual experimental phase, we introduce "scenario files," which describe structured tasks used in order to describe the expected, by either real user or VUM, activity during the computer-based experiment. Those files initially describe tasks in an abstract format (e.g., locate another user in the Metaverse), but using a scenario editor tool we developed that they are matched to actions performed on the actual user interface by annotating which areas of the testing interface correspond to active GUI elements in order to match the tasks with the related GUI elements. Eventually, the annotated interface is used in making connections between the interaction events and the tasks described in the scenario (simulation preparation). Figure 3 gives an overview of the simulation framework and the basic data flow between the main components.

During the simulation phase, initially we constructed three representative cognitive models: (a) elderly, (b) MCI, and (c) young users' group. These models were thereafter used to simulate scenarios similar to the ones performed by respective real users, toward determining and validating the comparison capacity of the system. Task validation was expanded in iterative trials with slightly different scenarios structured as sets of tasks. Those tasks were randomised so as to mitigate the learning effect. The results of the initial validated cognitive tasks were correlated and analysed with regard to the overall duration and the number of the low-level (LL) interaction events related to the input devices produced during the experiment.

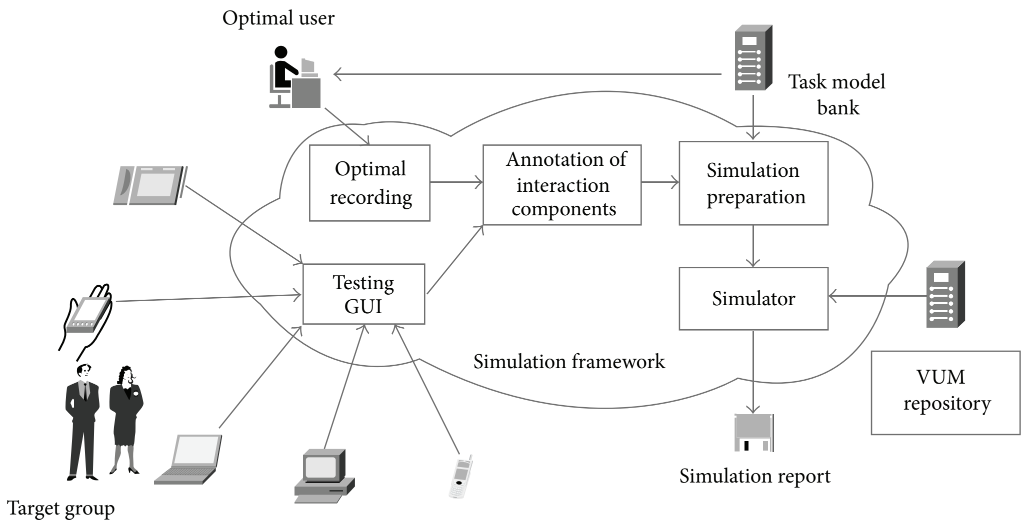


FIGURE 3: The basic components and data flow of the simulation framework.

The LL events included keystrokes cursor X-Y movement on the screen and mouse button clicks all hooked by a custom software tool from the lower OS layers (operating system messaging system). According to our simulation approach, the number of interaction events worked as an error rate indicator. Additionally, the "soft fail" concept was introduced to allow VUMs to retry the failed tasks in simulation time. Correspondingly, the "hard fail" concept terminates the simulation at the very first fail-point. In "soft fails" the VUMs have the chance to retry on failed tasks up to the limit of 200% of the OpUs time. This limitation allowed VUMs to repeat internal trial-and-error cycles without exceeding the doubled OpUs duration. We introduced this technique of VUMs to retry on failed tasks in order to prevent unsuccessfully completed tasks by a constrained VUM in infinite time (endless loop).

Finally, experimental results were exported in simulation reports which contained detailed information about the VUM's performance. Simulation reports were structured with respect to an XSD schema definition to formally describe measured elements as extensible markup language (XML) documents. A short human-readable description on the findings, including statistics on error rates, durations, number, and type of OS interaction events was included in the header of the simulation reports in order to allow results comparison with the performance of the real users (saved in log files). Real users performed in an external-to-the-simulator software environment, namely, the MCI screening tool and the real user's logging approach.

The accessibility assessment approach using VUMs proposed herein could effectively describe memory decline in mild cognitive impairment. The main quantitative measures were (a) the number of interaction events produced by the user during the scenario execution and (b) the time needed to complete the scenario (duration). Infotainment results were initially checked for correlation with age and MCI results, to construct a prediction model upon statistical methods using SPSS v.19. Moreover, the proposed cognitive models, based on the new dual cognitive module, were evaluated over whether they are better predictors of the time that actual MCI patients need to complete the tasks than the previously used cognitive architecture, without the new MCI submodule.

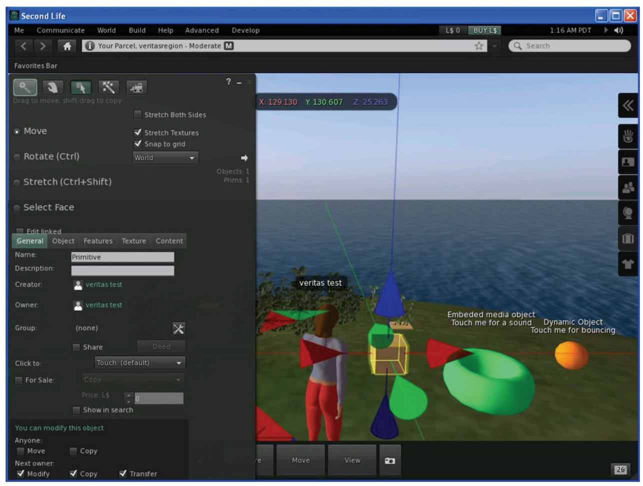


FIGURE 4: The Metaverse viewer used here as the testing interface.

4.2. Methods of Study 2. Pilot studies with elderly and MCI users took place on September 2013. The second life (SL) viewer of (Figure 4) was used to allow participants interact with the Metaverse and other distant visitors. The special version of the SL application used in the infotainment tests was developed using the OpensSim server platform and the source code version of the official SecondLife viewer application. This way, a private and fully controlled by the authors' team SL Metaverse environment—not visible from other SL visitors—was made available for subjects (real users and VUMs). It is worth noting that the capability to control the custom Metaverse refers not only to the features of the

Table 2: Scenarios performed in the infotainment pilot test.

id	Scenario name	Scenario description	Tasks	Required abilities
S1	Enter the Metaverse	It is required that users type username and password	5	Memory
S2	Change Outfit	Having a second outfit available, users are asked to change from outfit 1 to outfit 2	5	Decision-perception
S3	Upload file in Metaverse	Choose an image file from the local drive and upload	8	Information orientation
S4	Build 3D Object	Create a new cube in the virtual environment	5	Perception-reflection
S5	Scale 3D Object	Scale the cube to make its side equal to 1 m	5	Perception-reflection
S6	Rotate 3D Object	Rotate the cube in x, y, and z using the colored rotation rings	8	Motor-vision
S7	Move 3D Object	Move the 3D object in 3 directions using the colored moving arrows	8	Motor-vision
S8	Navigate Avatar in Free Mode	Rotate the head-camera of the avatar in space and then move few steps forward	3	Visual-motor
S9	Navigate Avatar to Sound Source	Navigate the avatar from a random position to the source of the sound	4	Visual-acoustic-motor-decision-perception
S10	Interact with Dynamic Object	Touch an object with dynamic behaviour	2	Motor-perception
S11	Interact with Multimedia Object	Touch an object which makes a sound	2	Motor-perception
S12	Initiate Chat with Another User	Locate another user in the Metaverse and send a "Hello" message	5	Perception-verbal
S13	Share Folder with Another User	Share folder with another user	9	Information orientation-memory

interface, but to the content of the immersive environment as well (2D and 3D content and functionality).

The infotainment tests were completed in the same day as the computer-based MCI screening test (Figure 4). In tests' iteration the scenarios were presented at a slightly different manner in order to eliminate the learning effect. However, the difficulty level of tasks has been maintained at both pilots in order to be comparable.

After performing the MCI screening tests, participants were entering a spacious demo room reserved for the pilot tests, usually in pairs. To ensure participants feel confident in using the Metaverse interface, after a demonstrating execution of all scenarios by the RP, the elderly participants were asked to follow. Thirteen relatively easy scenarios of common infotainment tasks were created for the test purposes. Some examples of scenarios are (Table 2): "enter the Metaverse" which is the user authentication process (S1), "change outfit in avatar's appearance" (S2), "create and manipulate 3D objects" (S4–S7), "navigate avatar" (S8-S9), "interact with objects" (S10-S11) and "interact with other visitors" (S12-S13). The length of each scenario, represented by the "number of tasks" indicates how many tasks must be completed by the user successfully to achieve the scenario goals.

4.3. Results of Study 2. The infotainment findings indicated that the number of interaction events, which also expresses the error rate due to the "hard error" and "soft error" concept, is a not a strong separator for the elderly and the MCI groups, so it was gradually repealed. Table 3 presents in detail the results of the infotainment pilot test organized per scenario for the two testing groups (elderly and MCI).

The research hypothesis here is "*if some portion of the population has cognitive impairments which affect memory, then for those who failed in the MCI screening test, it is expected that their scores in the related infotainment scenarios will be lower than for others.*" It is critical that not all the scenarios are equally interesting for the dual cognitive module development. For example, the scenarios where mostly motor, visual, and acoustic abilities are needed for their successful performance were intact for the MCI users. Indeed, among the available GUI scenarios, those which require more complex mental processes proved to be the appropriate for the distinction between MCI and healthy elderly users. Thus, in this study, emphasis is given on those scenarios which are primarily affected by memory and decision making decline.

Having the scores expressed as percentage differences of the OpUs performance, healthy elderly perform the infotainment scenarios in the 432.94% of the OpUs time and they produce the 50.77% of the OpUs events. Similarly, the results for the MCI users are 446.48% of the OpUs time and 56.20% of the OpUs events. Although the average values show that MCI users need more time and more keyboard and mouse clicks to perform the same scenarios; this extra effort may not be significant.

After normality tests, all variables related to scenario duration and events were tested for equality of means in the two testing groups (healthy elderly users and MCI users). For healthy elderly, the time needed to complete the first scenario (S1: $M = 59.53$, $\mathrm{SD} = 21.85$) and the last one ($M = 83.43$,

Table 3: Performed scenarios and scores.

Scenario	Optimal user		Healthy elderly user		MCI User	
	Events	Time*	Events	Time* (SD)	Events	Time* (SD)
S1	46	15.35	59.82 (23.43)	59.53 (21.85)	53.63 (5.07)	87.00 (32.34)
S2	10	8.73	13.64 (2.65)	46.88 (19.07)	15.00 (3.20)	56.75 (15.55)
S3	16	15.65	26.18 (9.44)	104.180 (44.35)	24.25 (8.37)	126.54 (44.07)
S4	10	6.64	11.91 (2.70)	35.37 (16.93)	13.38 (2.56)	47.33 (21.91)
S5	10	12.82	15.64 (7.31)	55.65 (32.80)	17.00 (3.38)	73.65 (23.19)
S6	12	12.15	20.36 (4.34)	63.3 (19.98)	20.88 (7.08)	63.50 (18.54)
S7	12	15.36	19.82 (4.99)	54.08 (19.75)	28.50 (16.41)	69.78 (25.901)
S8	6	4.59	7.82 (2.60)	14.07 (6.33)	6.88 (1.24)	16.93 (7.09)
S9	12	7.39	39.36 (25.31)	42.10 (16.24)	23.63 (9.70)	30.73 (15.55)
S10	4	1.51	5.09 (1.86)	10.10 (6.54)	5.38 (2.66)	12.36 (7.80)
S11	4	1.03	4.45 (0.82)	5.20 (3.80)	5.57 (4.20)	7.01 (7.42)
S12	22	12.79	26.64 (8.60)	51.97 (23.00)	26.13 (9.07)	47.68 (10.77)
S13	27	14.86	33.64 (10.14)	83.43 (25.05)	42.50 (23.39)	121.53 (45.42)

*Time is in seconds.

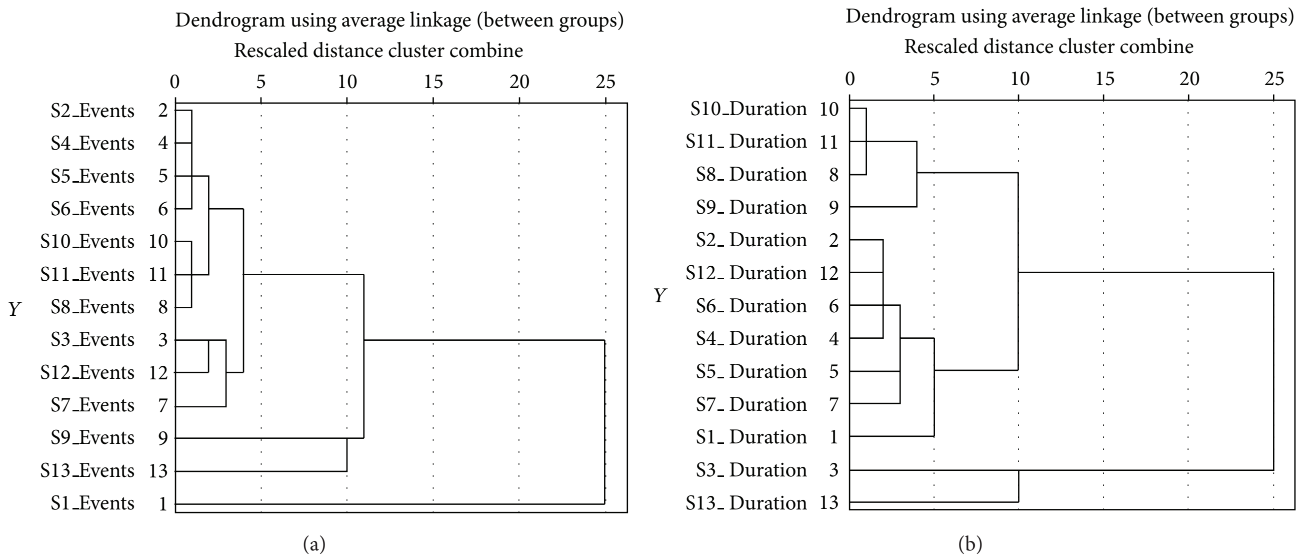


Figure 5: Two dendrograms as result examples which represent the variables of the number of interaction events (a) and time in seconds (b) using average linkage between groups.

SD = 25.05) was found to be significantly different than the duration of the first ($M = 87.00$, SD = 32.34) and the last scenarios ($M = 121.53$, SD = 45.42) of those having MCI (S1: $t = -2.21$, $P < 0.05$ and S13: $t = -2.34$, $P < 0.05$). Also, this difference was found at a significant effect size: $r = 0.445$ for S1 and $r = 0.461$ for S13. For all other scenario durations, the null hypothesis for equity of means was not rejected.

Regarding the other test metric, the number of interaction events (keystrokes and mouse clicks produced during the experiments), the null hypothesis for equity of means was not rejected ($P > 0.05$) in all tests. Concerning the number of interaction events, results indicate no difference between the two target groups. The same result was extracted from the test (Mann-Whitney U test) between the institutionalized and noninstitutionalized MCI.

4.4. Brief Discussion of Study 2 Results. The duration and the number of events for each scenario are not always following each other's fluctuations. Strong correlations between the duration and the number of interaction events were found only on scenarios S2 ($r = 0.730$, $P < 0.001$), S5 ($r = 0.669$, $P < 0.05$), S6 ($r = 0.525$, $P < 0.05$), S9 ($r = 0.477$, $P < 0.05$), and S11 ($r = 0.859$, $P < 0.001$). In Figure 5 two representative dendrograms of the duration and events variables are presented as the result of the hierarchical variable clustering. The vertical axis represents the number of interaction events (Figure 5(a)) and the overall duration (Figure 5(b)).

Following the results of the running data through the average linkage between groups, it is obvious that the two memory-related scenarios S1 and S13, as well as the two information orientation scenarios S3 and S13 are outliers and are fused in arbitrarily at much higher distances than others. The main aim was to extend all the knowledge acquired by the statistical analysis of the infotainment data in order to optimise the VUMs performance towards a second generation of VUMs, especially designed and tested for the selected MCI group.

The following section describes how the measures of the experiment were fitted to a regression equation in order to develop a performance prediction model for the second generation of VUMs. The expecting generation of VUMs, optimized based on the findings of study 2, summarizes all our expectations regarding the prediction of real MCI user's performance when working on interface designs.

5. Study 3: Optimization of VUMs

5.1. Introduction to Study 3. The currently presented effort is about the optimization of VUMs based on real users' data and the cognitive architecture proposed by the ACT-R model. The ultimate goal was to create robust VUM cognitive models as an internal mental representations of elderly users that will constitute their exact cognitive profile. In a first level the insertion of the MCI submodule as a perceptual module targeted at better estimations for early dementia prodromes. The new challenge was to determine the error rates and the extra time needed by MCI-VUMs and compare timing and success rate results with those generated by real MCI users. In a second level, a more detailed representation of the VUM cognitive model was used in interface accessibility assessment in a simulator. The results extracted from the infotainment test were used to compute which cognitive parameters should be updated.

5.2. Methods of Study 3. A regression equation was used to find out what relationship, if any, exists between the durations of performed—by real users—infotainment tasks and MCI screening test results given by the same subjects. As long as the predicted VUM's scores, as outcomes of the simulation tests, can fit to a regression equation, we can remodel the cognitive submodule of the existing VUMs for future experiments (optimized VUM).

This section explains the results of the forward selection procedure we followed as a strategy to choose the order of the approximate polynomial which calculates the scenario durations given the MCI screening scores. Using the BNT score as the independent variable, we successively fitted the models in increasing order and tested the significance of regression coefficients at each step of model fitting.

At first we assumed linearity as a starting point for calculating an equation which minimizes the distance between the fitted line and the experiment measure points (scenario durations). The linear regression equation was considered statistically significant ($P < 0.05$) and it gave a nice fit in our experimental results. Next, following our forward selection strategy, we upgraded our model to become a second-order model by trying to keep the statistical significance of the highest order term in the range of accepted values (<0.05) and at the same time to increase the R-squared (R^2) statistic (coefficient of determination). Calculated as the percentage of the response variable variation that is explained by our linear model, the R-squared statistic was found to be equal to 0.322. Keeping this value for future reference, the quadratic model can be expressed as

$$Y = X^2 b_2 + X b_1 + c, \tag{3}$$

where Y: duration of scenario, X: MCI screening score (BNT), and c, b_1, b_2 are calculated by the equation system.

And for n being the size of the data

$$\sum Y_i = nc + b_1 \sum X_i + b_2 \sum X_i^2, \tag{4}$$

$$\sum X_i Y_i = c \sum X_i + b_1 \sum X_i^2 + b_2 \sum X_i^3, \tag{5}$$

$$\sum X_i^2 Y_i = c \sum X_i^2 + b_1 \sum X_i^3 + b_2 \sum X_i^4. \tag{6}$$

The above method was used to calculate the expected duration of the scenarios when MCI users took part in the experiment. The final simulation report gives as predicted duration

$$D_i' = D_i + C_i', \tag{7}$$

where D_i': duration of scenario i for a VUM with 50% MCI, D_i: duration of scenario i, for a healthy elderly VUM, and C_i': constant coefficient.

As an example, given the duration of scenario 1 for a healthy elderly, the coefficients of (3) for calculating the duration of the scenario using a 50MCI VUM were found $b_1 = 41034.988$, $b_2 = -979.606$, and $c = -399951.739$.

In order to keep the simple rule that cognitive decline (e.g., MCI) adds extra time to the duration of scenarios, C_i' in (7) must be always greater than zero (nonnegative). So, the standard coefficient c, as calculated before, subtracts D_i from the result of (3) in order to express the amount of time added to the healthy elderly as a result of the cognitive decline.

The quadratic model gave an accepted statistical significance ($P < 0.05$) and moreover it offered a significantly higher R-square ($R^2 = 0.459$) than the linear model ($R^2 = 0.322$). By carrying on the model upgrading procedure, the next polynomial regression model (the cubic) was found not much better ($R^2 = 0.456$, $P < 0.05$) than the previous one and thus we did not upgraded our model more. The second-order model was more preferable due to its simplicity, so the quadratic effect parameter remained as the highest order term in our prediction model. To be noted is that other models, like the logarithmic, power, and exponential, were used to give no better results than the quadratic model.

The prediction models used for calculating the more accurate task durations of the MCI VUMs were incorporated into the GUI simulator to achieve optimized simulation results. Thus, the simulation reports produced on the later stages of the experiment contained more accurate results and were better fitted to the GUI designer's decision making. The optimized VUM behaviour was now both visually observed during simulation and imprinted in experiment reports. This was achieved by taking into account the updated parameter

Table 4: Test the new VUMs by example.

Users	Comments	S1 duration in sec	Fraction of the average MCI user
OpUs	Duration of the optimal user	15.35	0.17
Healthy Elderly	Score recorded in the infotainment pilot test	59.53	0.68
MCI	Score recorded by the actual MCI users	87.00	1.00
50Elderly	VUM of the first generation	17.30	0.19
50MCI	VUM of the second generation (MCI-ready)	78.38	0.90

values which were responsible for the control of the simulation time flow. The optimized time management regulation mechanism had a direct effect on the task completion times of the MCI VUMs, while in VUMs free of abnormal cognitive decline, the GUI simulation functionality remained unaffected. To be noted is that the optimized functionality of the MCI VUMs refers only to the time delays related to the cognitive decline and not the other kinds of time regulators (like in the motor submodule, e.g.).

The optimized version used the same motor, vision, and perception submodules as in any other case. Only when an MCI VUM was present in a simulation scenario task, the simulator adjusted its time management functionality according to the quadratic prediction model inserted into the optimized VUM. Practically, this was implemented as series of time delays between the task sequence calls. The rate and duration of these delays were corresponding to the degree of the cognitive decline (expressed as a "hypothetical BNT test" result). Finally, when the experimental simulation tasks were performed, the visually perceived difference between a healthy elderly and an MCI VUM (macroscopical difference) was like an unexpected or unjustified delay in the cursor motion planning. This distortion on cursor movement was due to the simulated "forgetfulness" of the VUMs which had lost control over their performed tasks, although their tasks were simply noted in the loaded scenario file.

Another important time regulator of the optimized model is the time needed by the VUM to select one option in the interface instead of another, when many options were present at the same time. This is quite similar to the common question "Which button does the job, this one or some other?" The simulator takes into account the Hick-Hyman law as already described, but not the human processing speed that may be affected because of the intelligence quotient (IQ) of the VUMs [36]. Real users were not asked to give an IQ test before or after the test and so the IQ was not taken into account by the optimized VUM model.

The outcome to this point is a more detailed representation of the VUM cognitive model for use in interface accessibility assessment in a simulator. The results extracted from the infotainment test were used to compute which cognitive parameters should be updated and by which factor in order to eliminate the differences between the scores of VUMs and the real elderly users.

5.3. Results of Study 3. The new cognitive architecture, inserted into the cognitive-aware VUM was evaluated against the old VUM generation to measure the benefit from the adoption of the new cognitive model. As an example, in Table 4, a comparison between the time scores of the real MCI users, the 50Elderly (old VUM) and the 50MCI (new VUM) in seconds are presented. Data were extracted by simulating the "Enter the Metaverse" scenario (S1) and the optimal user's performance has been subtracted.

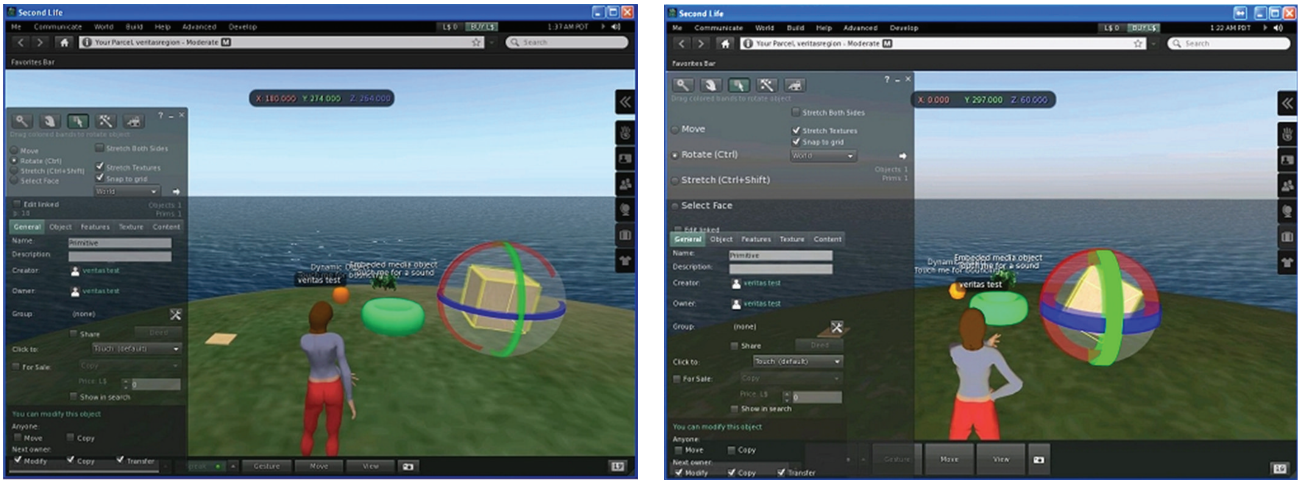
Practically, the optimization of the cognitive models was proved by using statistical methods over a set of experimental time and score recordings. A number of regression models have been used for prediction purposes as it was previously described. The lower and upper bounds of the 95% confidence interval for each of the coefficients of the regression equation were used to produce the range of the percentage MCI VUMs, starting from the barely 1MCI up to the maximum 100MCI. As an example when the OpUs needs 15.35 sec to complete the user authentication scenario (S1), the healthy elderly need 59.53 sec on average as seen in the infotainment pilot test, and the average MCI user needs 87 sec for the same task.

The 50Elderly VUM needed 17.30 sec, which is far away from the actual user because a VUM does not take into account common typing mistakes made by the average user or momentary lapse of memory (e.g., forgetting one's password). This exemplary performance is not convincing because it is closer to the OpUs than to the elderly or the MCI users. The only time-penalty paid above the OpUs duration is an age-related delay inserted in the motor module (cursor moves) to justify the difference (1.95 sec).

The 50MCI on the other hand, created with BNT = 22 in cognitive scores, needed 88.21 sec that is 98% similar to the actual MCI user's score, in contrast with the old 50Elderly VUM which needed only the 19% of the actual time score.

5.4. Brief Discussion of Study 3 Results. The purpose of this study was the optimization of VUMs performance that was made possible after a set of testing scenarios, on actual elderly and MCI users, to fit better the performance of the real persons. The infotainment application area was selected as representative of typical computer-based environments, encompassing from 2D GUI elements to immersive interactivity and social communication with other distant users.

The duration of the scenarios which require strong memory (remember username and password in S1, the sequence of the dialog boxes to open in S13) was found to be affected



(a) (b)

Figure 6: The original Metaverse interface (a) and the new design (b).

by the existence of MCI. The same can be generalized for "deep GUIs" which require that users follow long point and click routes to master the application, even if the application itself has a good learning curve. The number of interaction events produced during testing, which hides the success and failure rates of mouse clicks and keyboard strokes, was proven a weak predictor, probably because of the scenario directions themselves: the fact that the elderly and MCI testers were given directions of how to perform the scenarios resulted in inadvertently correcting errors related to the decision-making. A nonsupervised pilot test with real users based on a set of scenarios clearly targeted to the memory-judgment dipole, in which only the goal is shared with the participants, would be a more sensitive instrument to detect time delays and interaction patterns in MCI populations.

Following a deterministic approach, the VUM scores of every previous test are used to predict more accurately the behavioural and cognitive alterations of actual end users by time. This approach allows interface designers to run accessibility tests on their new designs at early development phases without the need to recruit real impaired persons.

The new cognitive module, the development of which was based on the total scenario duration prediction model (quadratic model), worked as expected and produced timing scores closer to the actual MCI users than the previous model. Bearing in mind that the achieved quality in the prediction model development is characterized by the coefficient of determination R^2, in the field of human behaviour, such as the infotainment experiment, it is entirely expected that R-square values will be lower than 25% [37, 38]. The performance of human subjects in complex tasks over complex user interfaces is simply harder to predict than in other research fields. Thus, the resulting value of 45.9% for R-square satisfies our research expectations. From the VUM optimization point of view, this was a key-point result because it let us draw the important conclusion that the MCI scores of BNT can explain almost the half of the variability of the response data around its mean.

6. Study 4: Accessibility Assessment with Optimized VUMs

6.1. Introduction to Study 4. This simulation test using the new VUMs based upon the optimized GUI aimed at supporting more tangible graphical user interfaces and directed the Metaverse developers to redesign the user interface according to the needs of all potential users, including those having MCI (Figure 6). Towards VUM optimization according to the needs of MCI users, some dialog boxes have been simplified enough to minimize the cognitive load, while pictograms were preferred for triggering common tasks and thus eliminating memory retrieval.

6.2. Methods of Study 4. The new interface was tested by a new group of VUMs, starting from 50MCI and reaching 85MCI. Instead of creating a completely new VUM, the elderly VUM was modified in such a way that an architectural extension (MCI submodule) is affecting its performance in the same way MCI is affecting humans. The duration predictions of the optimized VUM were sliding in the range indicated by the lower and upper bounds of the model coefficients as expected. Above the 90MCI threshold, the interface was considered by its creators as not capable of being redesigned without losing its basic functionality.

6.3. Results of Study 4. The test performed following the same protocol on the updated interface yielded better results, namely, reduced times needed to complete the scenarios in average by 14.09%. In the memory-demanding scenarios (S1 and S13), which are of particular interest for MCI users, the elderly reduced the user authentication time to 57 sec (SD = 23.36) from 59 sec (SD = 21.85), but MCI users reduced to 65 sec (SD = 34) from 87 sec (SD = 31.54).

6.4. Brief Discussion of Study 4 Results. Accessibility assessment through the optimized VUMs has practically proved that inserting the novel MCI cognitive submodule in the overall VUM architecture provides the potential that interface designers can make tests on the simulator for decision making on interface redesign. The results were proven valuable as the new accessible interface design reduced the memory load in memory demanding tasks and also the cognitive load in decision making tasks. The new generation of VUMs, the cognition-enabled ones, responded in a human-like way pointing to the interface designers which design feature changes were positives according to the accessibility rules.

7. General Discussion of All Studies

The simulator used to perform experiments through VUMs was developed to be results-oriented. The ultimate goal was to enhance interfaces' quality in order to support designers performing realistic simulations. On the other hand, the ACT-R model, in which the structure of our VUMs was based, was initially proposed as a meaning-oriented structure for general user modeling use. The multiple studies and experiments we performed targeted the development of a novel cognitive model customized to MCI users as an extension to existing cognitive models.

Each study has contributed in its own way in the overall effort to simulate dementia for interface accessibility assessment. The first study provided the necessary quantitative data to describe the medical state of the participants on the basis of well-established MCI screening tests. The second study has served as a systematic semiautomated observation of MCI and controls over a period of time, during task oriented scenarios. Among the performed tasks, those pointing on memory and decision making processes were of particular importance for the target audience of patients suffering from dementia. The third study indicated the relation between the performed timing scores and the state of cognition. From the results we gathered, the model which best described in statistical importance the relation of user's performance and cognitive decline associated with MCI was a quadratic model. Using these outcomes, we optimized our models according to the results of the previous study checking also the consistency with the simulator requirements. The cognitive submodule in particular, which was inserted into the existing virtual user architecture introducing a new generation of cognitive-aware VUMs, is capable of performing the same tasks as real users in a more realistic manner. The same testbed (Metaverse environment) was used to evaluate the expected benefits on VUM's quality, as a matter of prediction ability. Indeed, the novel VUMs achieved simulation performances closer to those of the real users. With a better VUM quality available, the last study was performed on a real world interface problem concerning the interface accessibility assessment for GUI designers. The Metaverse designers redesigned its interface so as not to exclude MCI users from the potential body of target users. The new design was evaluated by the novel VUMs to indicate results according to which, when the MCI cognitive submodule was enabled and active, the simulator produced results closer to the real MCI users, especially in memory-demanding and decision making tasks.

The followed user-centric methodology and mainly the novel VUMs created by this study can be used in most cases of interface accessibility assessment covering over 90% of all common interface accessibility issues. It is limited only by an extreme cognitive disability threshold which corresponds to the 90% MCI users. VUMs of very high dementia levels can produce valid simulation results, but their duration predictions cannot always be valuable for interface designers as extreme interventions on an interface design may cause the interface to lose an important part of its functionality.

8. Conclusion

Commonly, GUI designers rescale GUI elements or increase color contrast to serve the needs of elderly or impaired users. However, there is still need for specific interface design strategy for the reconfiguration and change of visual metaphors to customize interfaces to the specific needs of groups with cognitive related limitations. Our approach combined literature findings and experimental data to predict in better accuracy the performance of cognitive impaired VUMs over known interface accessibility limitations. In the proposed VUM architecture implementation, we inserted an additional cognitive submodule to produce "noisy" functionality in the internal VUM structure instead of producing noisy data in the output.

Pilot tests on the simulator indicated high correlation between the tests scores of real users and VUMs. These findings should be viewed as an early approach about the latent interest in encompassing cognitive features in the design of infotainment simulation models. Evaluation tests from infotainment products have shown that the proposed VUMs' modification customized according to cognitive impairments can be a valuable asset to the designer, as a way of increasing products' and services accessibility for elderly with varying deficits of memory, attention, judgment, and communication ability. The utilised scenarios required not only memory but even more complex cognitive functions, such as decision making abilities.

After the encouraging results extracted from the pilot test presented herein, the forthcoming tests will be performed by recruiting a wider body of subjects (30–50 persons). Those tests will include the test-retest reliability evaluation of the computer-based screening test and, moreover, it is expected to improve the prediction accuracy of the simulation results between VUMs and actual users. Recruitment criteria could be extended so as to include a depression detection test like the Beck Depression Inventory (BDI-II) to eliminate possible effects caused by extraneous to the MCI itself factors. Moreover, among our future plans is the actual use and evaluation of vocal and aural modules.

This research effort, being in-line with the "design for all" principle, is helping people with cognitive functionality decline to improve their quality of life in an indirect way: not by providing therapeutic appliances straight to them, but with changing the design culture of the industrial world on their behalf.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work has received funding from the European Union's Horizon 2020 research and innovation program under Grant agreement no. 643433, project RAMCIP.

References

[1] S. J. Czaja and C. C. Lee, "The impact of aging on access to technology," *Universal Access in the Information Society*, vol. 5, no. 4, pp. 341–349, 2007.

[2] D. G. Turnbull, C. M. Chewar, and D. S. McCrickard, "Are cognitive architectures mature enough to evaluate notification systems," in *Proceedings of the International Conference on Software Engineering Research and Practice (SERP '03)*, Las Vegas, Nev, USA, June 2003.

[3] D. Sloan, M. T. Atkinson, C. Machin, and Y. Li, "The potential of adaptive interfaces as an accessibility aid for older web users," in *Proceedings of the International Cross Disciplinary Conference on Web Accessibility (W4A '10)*, ACM, April 2010.

[4] C. DeCarli, B. L. Miller, G. E. Swan, T. Reed, P. A. Wolf, and D. Carmelli, "Cerebrovascular and brain morphologic correlates of mild cognitive impairment in the national heart, lung, and blood institute twin study," *Archives of Neurology*, vol. 58, no. 4, pp. 643–647, 2001.

[5] N. Taatgen and J. R. Anderson, "The past, present, and future of cognitive architectures," *Topics in Cognitive Science*, vol. 2, no. 4, pp. 693–704, 2010.

[6] F. E. Ritter and R. M. Young, "Embodied models as simulated users: introduction to this special issue on using cognitive models to improve interface design," *International Journal of Human Computer Studies*, vol. 55, no. 1, pp. 1–14, 2001.

[7] R. Sun, "Introduction to computational modeling," in *The Cambridge Handbook of Computational Psychology*, R. Sun, Ed., pp. 3–19, Cambridge University Press, Cambridge, Mass, USA, 2008.

[8] J. R. Anderson and C. Lebiere, *The Atomic Components of Thought*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1998.

[9] D. E. Kieras and D. E. Meyer, "An overview of the EPIC architecture for cognition and performance with application to human-computer interaction," *Human-Computer Interaction*, vol. 12, no. 4, pp. 391–438, 1997.

[10] B. E. John and D. E. Kieras, "The GOMS family of user interface analysis techniques: comparison and contrast," *ACM Transactions on Computer-Human Interaction*, vol. 3, no. 4, pp. 320–351, 1996.

[11] P. M. Langdon, M. F. Gonzalez, and P. Biswas, "Designing studies for the requirements and modelling of users for an accessible set-top box," in *Proceedings of the 8th International Conference on Disability, Virtual Reality & Associated Technologies*, pp. 203–212, Valparaiso, Chile, 2010.

[12] J. Mankoff, H. Fait, and R. Juang, "Evaluating accessibility by simulating the experiences of users with vision or motor impairments," *IBM Systems Journal*, vol. 44, no. 3, pp. 505–517, 2005.

[13] M. Sutterer, O. Droegehorn, and K. David, "UPOS: user profile ontology with situation-dependent preferences support," in *Proceedings of the 1st International Conference on Advances in Computer-Human Interaction (ACHI '08)*, pp. 230–235, IEEE, February 2008.

[14] G. Trafton, L. Hiatt, A. Harrison, F. Tanborello, S. Khemlani, and A. Schultz, "ACT-R/E: an embodied cognitive architecture for human-robot interaction," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 30–55, 2013.

[15] M. D. Byrne, "ACT-R/PM and menu selection: applying a cognitive architecture to HCI," *International Journal of Human-Computer Studies*, vol. 55, no. 1, pp. 41–84, 2001.

[16] T. J. Wong, E. T. Cokely, and L. J. Schooler, "An online database of ACT-R parameters: towards a transparent community-based approach to model development," in *Proceedings of the 10th International Conference on Cognitive Modeling (ICCM '10)*, pp. 282–286, August 2010.

[17] D. D. Salvucci, E. R. Boer, and A. Liu, "Toward an integrated model of driver behavior in cognitive architecture," *Transportation Research Record*, vol. 1779, no. 1, pp. 9–16, 2001.

[18] L. Razmerita, A. Angehrn, and T. Nabeth, "On the role of user models and user modeling in Knowledge Management Systems," in *Proceedings of the International Conference on Human-Computer Interaction (HCI '03)*, pp. 450–456, 2003.

[19] D. Heckmann and A. Krueger, "A user modeling markup language (UserML) for ubiquitous computing," in *User Modeling 2003: Proceedings of the 9th International Conference on User Modeling (UM '03), Johnstown, Pa, USA, June 22–26*, vol. 2702 of *Lecture Notes in Computer Science*, pp. 393–397, Springer, Berlin, Germany, 2003.

[20] A. Serna, H. Pigot, and V. Rialle, "Modeling the progression of Alzheimer's disease for cognitive assistance in smart homes," *User Modelling and User-Adapted Interaction*, vol. 17, no. 4, pp. 415–438, 2007.

[21] P. Biswas, P. Robinson, and P. Langdon, "Designing inclusive interfaces through user modeling and simulation," *International Journal of Human-Computer Interaction*, vol. 28, no. 1, pp. 1–33, 2012.

[22] R. Bellamy, B. John, and S. Kogan, "Deploying CogTool: integrating quantitative usability assessment into real-world software development," in *Proceedings of the 33rd International Conference on Software Engineering (ICSE '11)*, pp. 691–700, ACM, May 2011.

[23] K. E. Nutter-Upham, A. J. Saykin, L. A. Rabin et al., "Verbal fluency performance in amnestic MCI and older adults with cognitive complaints," *Archives of Clinical Neuropsychology*, vol. 23, no. 3, pp. 229–241, 2008.

[24] N. Kaklanis, P. Moschonas, K. Moustakas, and D. Tzovaras, "Virtual user models for the elderly and disabled for automatic simulated accessibility and ergonomy evaluation of designs," *Universal Access in the Information Society*, vol. 12, no. 4, pp. 403–425, 2013.

[25] N. Kaklanis, P. Biswas, Y. Mohamad et al., "Towards standardisation of user models for simulation and adaptation purposes," *Universal Access in the Information Society*, 2014.

[26] J. Vanderdonckt, "A MDA-compliant environment for developing user interfaces of information systems," in *Proceedings of the 17th International Conference on Advanced Information Systems Engineering (CAiSE '05)*, pp. 16–31, June 2005.

[27] M. Beckmann, U. Yilmaz, G. Pöhler, and A. Wegerich, "A framework for task accomplishment using an ACT-R simulation," in

Proceedings of the 11th International Conference on Cognitive Modeling (ICCM '12), pp. 127–128, Berlin, Germany, April 2012.

[28] W. E. Hick, "On the rate of gain of information," *Quarterly Journal of Experimental Psychology*, vol. 4, no. 1, pp. 11–26, 1952.

[29] R. Hyman, "Stimulus information as a determinant of reaction time," *Journal of Experimental Psychology*, vol. 45, no. 3, pp. 188–196, 1953.

[30] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of Experimental Psychology*, vol. 47, no. 6, pp. 381–391, 1954.

[31] P. Moschonas, A. Tsakiris, I. Paliokas, and D. Tzovaras, "User interfaces accessibility assessment using virtual user models," in *Proceedings of the International Workshop on Personalisable Media Systems & Smart Accessibility*, Istanbul, Turkey, 2012.

[32] B. Chikhaoui and H. Pigot, "Analytical model based evaluation of human machine interfaces using cognitive modeling," *International Journal of Information Technology*, vol. 4, no. 4, pp. 252–261, 2008.

[33] Z. S. Nasreddine, N. A. Phillips, V. Bédirian et al., "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment," *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.

[34] E. Kaplan, H. Goodglass, and S. Weintraub, *Boston Naming Test*, Lea & Febiger, Philadelphia, Pa, USA, 1983.

[35] L. T. Connor, A. Spiro III, L. K. Obler, and M. L. Albert, "Change in object naming ability during adulthood," *Journals of Gerontology—Series B Psychological Sciences and Social Sciences*, vol. 59, no. 5, pp. P203–P209, 2004.

[36] E. Roth, "Die Geschwindigkeit der Verarbeitung von Information und ihr Zusammenhang mit Intelligenz/The speed of information processing and ist relation to intelligence," *Zeitschrift für Experimentelle und Angewandte Psychologie*, vol. 11, no. 4, pp. 616–622, 1964.

[37] M. L. Mitchell and J. M. Jolley, *Research Design Explained*, Wadsworth Publishing, Cengage Learning, 8th edition, 2012.

[38] J. Lawrance, R. Bellamy, and M. Burnett, "Scents in programs: does information foraging theory apply to program maintenance?" in *Proceedings of the IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC '07)*, pp. 15–22, September 2007.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 984353, 8 pages
http://dx.doi.org/10.1155/2015/984353

Review Article

Thermal Infrared Imaging-Based Computational Psychophysiology for Psychometrics

Daniela Cardone, Paola Pinti, and Arcangelo Merla

Infrared Imaging Lab, Institute for Advanced Biomedical Technology (ITAB), Department of Neuroscience, Imaging and Clinical Sciences, University of Chieti-Pescara, Via Luigi Polacchi 11, 66013 Chieti, Italy

Correspondence should be addressed to Arcangelo Merla; a.merla@itab.unich.it

Received 15 October 2014; Revised 5 January 2015; Accepted 27 January 2015

Academic Editor: Dimitris Giakoumis

Thermal infrared imaging has been proposed as a potential system for the computational assessment of human autonomic nervous activity and psychophysiological states in a contactless and noninvasive way. Through bioheat modeling of facial thermal imagery, several vital signs can be extracted, including localized blood perfusion, cardiac pulse, breath rate, and sudomotor response, since all these parameters impact the cutaneous temperature. The obtained physiological information could then be used to draw inferences about a variety of psychophysiological or affective states, as proved by the increasing number of psychophysiological studies using thermal infrared imaging. This paper presents therefore a review of the principal achievements of thermal infrared imaging in computational physiology with regard to its capability of monitoring psychophysiological activity.

1. Introduction

Understanding affective and psychophysiological states of a conversational interlocutor is fundamental for setting a proper communication, establishing social and affective ties, choosing social strategies, and maintaining a contingent interaction. Such understanding and the quantitative assessment of psychophysiological states represent one of the major challenges in applied psychophysiology and, more recently, one of the major issues in human-machine or human-artificial agent interaction.

In fact, a common key requirement for all typologies of the human-artificial agent interaction is to set up a contingent interaction, with the agent being capable of not only reacting to human actions, but also (or should) reacting in ways that are congruent with the emotional and psychophysiological state of the human user or interlocutor [1, 2].

Conventional approaches for assessing affective and psychophysiological states are based on the measurements of several physiological parameters expressing autonomic nervous system (ANS) activity, like skin sympathetic response (SSR), hand palm temperature, heart rate and/or breath modulations, peripheral vascular tone, facial expression, posture, gaze, and electromyography activity [3–5]. Apart from the assessment of facial expression, monitoring these parameters usually requires the use of contact sensors attached to the subject. More recently, some of them are monitored through watch-like or wireless devices.

In order to exceed limitations derived from the use of contact sensors, computational psychophysiology based on imaging approach can be recommended.

To this goal, thermal infrared (IR) imaging has been proposed as a potential solution for noninvasive and ecological recording of ANS activity [6]. Thermal imaging allows the contactless and noninvasive recording of the cutaneous temperature through the measurement of the spontaneous thermal irradiation of the body [7]. The psychophysiological activity can thus be assessed through its thermal effects recorded by thermal IR imaging. In fact, skin temperature is modulated by the ANS activity, which in turn regulates the cutaneous blood perfusion, the local tissue metabolism, and the sudomotor response [8–17]. Since the face is naturally exposed to social communication and interaction, thermal imaging for psychophysiology is generally performed by imaging the subject's face. Given the proper choice of IR imaging systems, optics, and solutions for tracking the regions of interest, it is possible to avoid any behavioral restriction of the subject [18, 19]. This possibility is particularly important, for example, in the developmental psychology or human-artificial agent interaction fields.

This paper reviews the state of the art in the field of thermal IR imaging-based computational physiology. The general intent of the paper is to provide insights about its potentialities and limits for its use in quantitative psychophysiology.

2. Thermal Signatures of Psychophysiological Signals

Thermal signatures of a variety of psychophysiological signals have been identified. In particular, it has been demonstrated that, through bioheat transfer models, it is possible to compute at a distance the cardiac pulse, the breathing rate, the cutaneous blood perfusion rate, and the sudomotor response. This section summarizes the methods and the results for computational physiology based on thermal IR imaging.

2.1. Cardiac Pulse. Thermal IR imaging allows the computation of the cardiac pulse at a distance through the modeling of the pulsatile propagation of blood in the circulatory system [9, 20–23]. In fact, the heart contraction during the ventricular systole generates a pressure wave, which propagates through the arterial tree. The arterial pulse reflects the heart activity thus providing a measure of cardiac interbeat intervals, heart rate, and its variability [22]. The method presented by Garbey and colleagues [9] is based on the hypothesis that the temperature modulation due to pulsating blood flow produces the strongest variation on the temperature signal of a superficial vessel. The proposed model simulates the heat diffusion process on the skin originating from the core tissue and a major superficial blood vessel. They took into account noise effects due to the environment and instability in blood flow. Their simulation demonstrated that the skin temperature waveform is directly analogous to the pulse waveform, except for its smoothed, shifted, and noisy shape because of the diffusion process. The method proposed by Garbey and colleagues [9] for computing heart rate is based on the information contained in the thermal signal emitted from major superficial vessels and recorded through a highly sensitive thermal imaging system. To compute the frequency of modulation (pulse), the authors extract a line-based region along the vessel. Then, they apply fast Fourier transform (FFT) to individual points along this line of interest, to capitalize on the pulse's thermal propagation effect. Finally, they use an adaptive estimation function on the average FFT outcome to quantify the pulse (Figure 1). Experiments on a data set of 34 subjects compared the pulse computed from the thermal signal analysis method to concomitant ground-truth measurements obtained through a standard contact sensor (piezoelectric transducer). The performance of the method ranges from 88.52% to 90.33% depending on the clarity of the vessel's thermal imprint. Sun et al. [20] applied the same method but working at 90 degrees across the direction of the target vessel. An extension of the abovementioned methods has been realized by Bourlai et al. [21]. They applied these two methods on an automatic tracked region of interest (ROI) and added noise reduction through a two-stage algorithm that discards problematic frames as a result of bad tracking. The new method was tested on 12 subjects and reduced the instantaneous measurement error from 10.5% to 7.8%, while it improved mean accuracy from 88.6% to 95.3%.

More recently, Farag et al. [22, 23] proposed an automatic method to determine arterial pulse waveforms through the use of thermal imaging. This method is based on the hypothesis of the quasiperiodic thermal pattern on the skin due to the arterial pulse to automatically detect the areas surrounding superficial arteries. Multiscale decomposition models, such as wavelet decomposition, are applied to each thermal image to extract those scales containing most of the arterial pulse information. The influence of irrelevant noise is thus minimized and the arterial waveform recovery is more accurate. The more coarse scales are used to track the region of interest (ROI). The finer scales are used to compute the arterial pulse through the periodicity detection (PD) algorithm: a region of measurement (ROM) is chosen within each ROI and different ROM configurations are tested (size, orientation, scale, and location); for each tested ROM, continuous wavelet analysis is run to remove high frequency noise and to extract arterial pulses structures; maxima are calculated from the resulting waveform which in turn correspond to the systolic peaks (used to compute heart rate, beat to beat, and heart rate variability). The PD algorithm individuates the optimal ROM in terms of the periodicity of the waveform and of its relevance to the true arterial pulse propagation. Validation of the method on 8 subjects showed perfect matching with oximeter data [23].

2.2. Breathing Rate. Breathing consists of inspiration and expiration cycles during which heat exchanges occur between airflows and nostrils. These exchanges create a periodic or quasiperiodic thermal signal in the proximity of the nostrils that oscillates between high (expiration) and low (inspiration) values (Figure 2). Thermal imaging can capture this phenomenon at a distance, achieving an accuracy of 96.43% [8].

In conventional respiratory studies, a thermistor is usually positioned near the nostrils to capture this phenomenon and produce a representative breath signal [24].

Thermal imaging acts therefore as a virtual thermistor, since it captures the same process, but at a distance. The estimation of breathing rate through thermal imaging is very accurate as proved by comparison with respiratory signals taken with conventional sensors [25, 26]. From the work of Murthy et al. [25], a high degree of chance-corrected agreement ($\kappa = 0.92$) was found between the airflow monitored through thermal imaging and oronasal thermistors. Correlation coefficients between the thermally and mechanically (LifeShirt technology; see [26]) recorded breath rate signals resulted as high as 1 over a sample of 25 subjects, in both shallow, normal, and forced ventilations [26].

Lewis et al. [26] showed also the possibility of estimating the relative tidal volume from thermal imaging. The

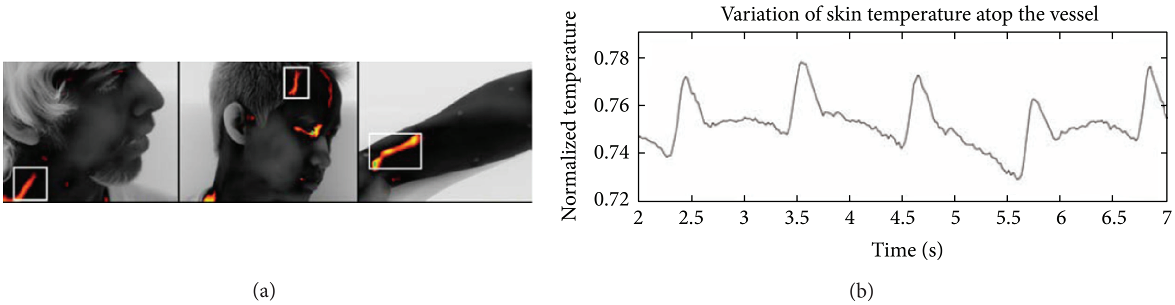


Figure 1: Pulse computation from thermal imaging data. (a) Collection point on the carotid arteriovenous complex, the frontotemporal region, and the wrist of the subject. (b) Temperature profile after removing frequency signals lower than 0.67 Hz (40 bmp) and higher than 1.67 Hz (100 bmp) (adapted from [9]).

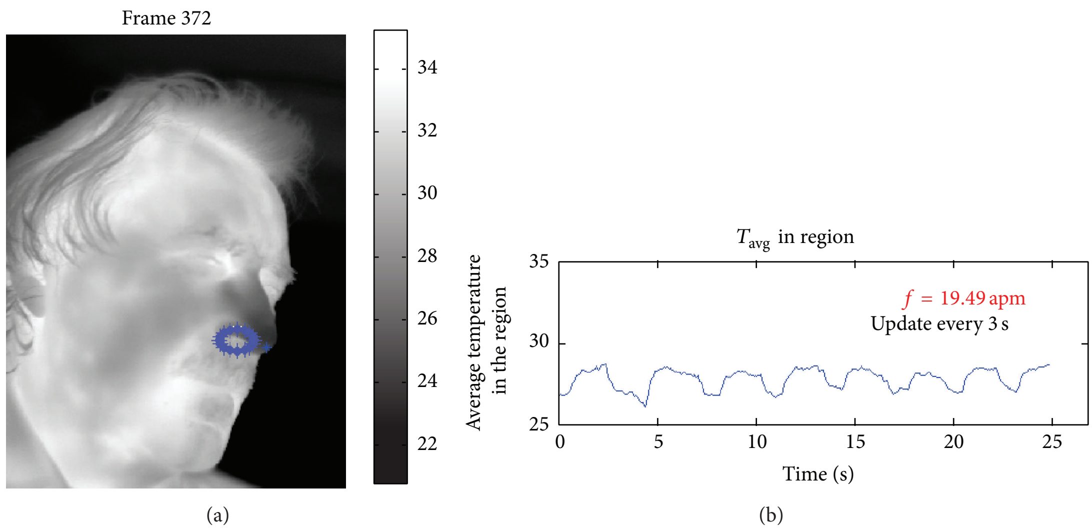


Figure 2: Thermal imaging data. (a) Thermal image showing the thermal track of the airflow. (b) Raw temperature versus time profile for a region of interest close to the nose tip.

correlation coefficient between the thermal and meccanical recordings over the same sample was 0.90.

Statistical methods have also been proposed to compute the contactless breathing signature. The algorithm used by Murthy et al. [27] is based on the method of moments and Jeffrey's divergence measure. This method has been tested on 10 subjects leading to a mean accuracy of 92% compared with the respiratory belt data at the thorax.

Multiresolution analysis has been used as well [28, 29]. Fei and Pavlidis [29] extracted the breathing content from the mean temperature of the nostrils through wavelet analysis. They found a high degree of agreement between the thermally recovered breathing waveform and the corresponding thermistor one in 20 subjects. In the work of Chekmenev et al. [28] the nasal region is tracked over time and for each frame the ROI is decomposed and averaged at three different scales. Wavelet transform is then applied to the resulting signal. The scale that contains most of the breathing information is extracted and used to compute the breathing rate. This approach has been tested on 4 subjects and the results perfectly matched with the piezoelectric measure device signals.

Thermal IR imaging has been also used to retrieve breath-related thermal variations from nasal, ribcage, and abdomen regions of interest in children, both healthy and with respiratory pathology. The study proved that thermal IR imaging reliably acquires time-aligned nasal airflow and thoracoabdominal motion without relying on attached sensor performance and detects asynchronous breathing in pediatric patients [30].

Fei and colleagues [31] proposed a novel methodology to monitor sleep apnea through thermal imaging. The nostril region was segmented and tracked over time via a network of cooperating probabilistic trackers. Then, the mean thermal signal of the nostril region, carrying the breathing information, was analyzed through wavelet decomposition. The experimental set included 22 subjects (12 men and 10 women). The sleep-disordered incidents were detected by

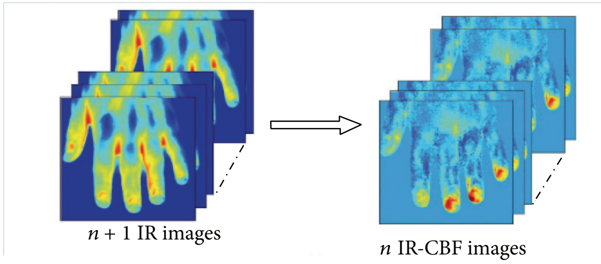


Figure 3: From the thermal IR image series to the cutaneous blood flow (CBF) images derived from thermal IR imagery. The series of IR images is converted into a series of IR-CBF images by applying computational models for bioheat exchange (adapted from [32]).

both thermal and standard polysomnographic methodologies. The high accuracy achieved confirmed the validity of the proposed approach for nonobtrusive clinical monitoring of sleep disorders [31].

2.3. Cutaneous Blood Perfusion Rate. Bioheat transfer models permit the calculation of the cutaneous perfusion from high-resolution IR image series (Figure 3) [32, 33]. Pavlidis and Levine [33] even suggested to use cutaneous perfusion rate changes in the periorbital region as a performing channel for a new generation of deception detection systems, based on the flight-fight response of the inquired subject to sensitive questions. The models adopted are derived from previous works of Fujimasa et al. [34], Pavlidis and Levine [33], and Merla and colleagues [32]. According to these models, cutaneous temperature change over a short time is mainly due to the heat gain/loss via convection attributable to blood flow of subcutaneous blood vessels and the heat conducted by subcutaneous tissue.

The models show that the blood flow rate and the cutaneous blood flow depend mostly on the time-derivative of the cutaneous temperature and on the difference between the temperatures of the cutaneous layers and the inner tissues [32].

It has been demonstrated that it is therefore possible to transform raw thermal image series in cutaneous blood flow image series (Figure 3).

The method has been validated by comparison with laser Doppler imagery (Figure 4). Merla and colleagues showed that, in twenty healthy subjects, cutaneous blood flow values, simultaneously computed by thermal IR imagery and measured by laser Doppler imaging, linearly correlate ($R = 0.85$, Pearson Product Moment Correlation) [32]. The method has been applied in psychophysiology for deception detection [35] and emotion assessment [10].

2.4. Sudomotor Response. Electrodermal responses have been among the most widely employed psychophysiological measures of autonomic nervous system activity. The Skin Conductance Response (SCR) and related measures, like galvanic skin response (GSR), have been shown to correlate with the number of active sweat glands, which activation can be easily visualized through facial thermal IR imaging by the appearance of cold dots over the thermal distribution of the face (Figure 5).

Concurrently to the palm area, strong sweat gland activation is manifested in the maxillary, perioral, and nose tip regions (Figure 5). Multiresolution analysis of the temperature changes reveals tonic (baseline and/or general) and phasic (event-related) components strongly correlated with GSR sympathetic constituents [12, 13, 16, 36]. For example, Pavlidis et al. [13] reported very high correlation coefficients between the GSR and the thermal measurement on the finger ($r_{\mathrm{MIN}} = 0.968$) and on the perinasal region ($r_{\mathrm{MIN}} = 0.943$). Moreover, wavelet analysis of thermal signals [12] revealed that the maxillary channel contains information about the sympathetic response almost as much as the GSR channel.

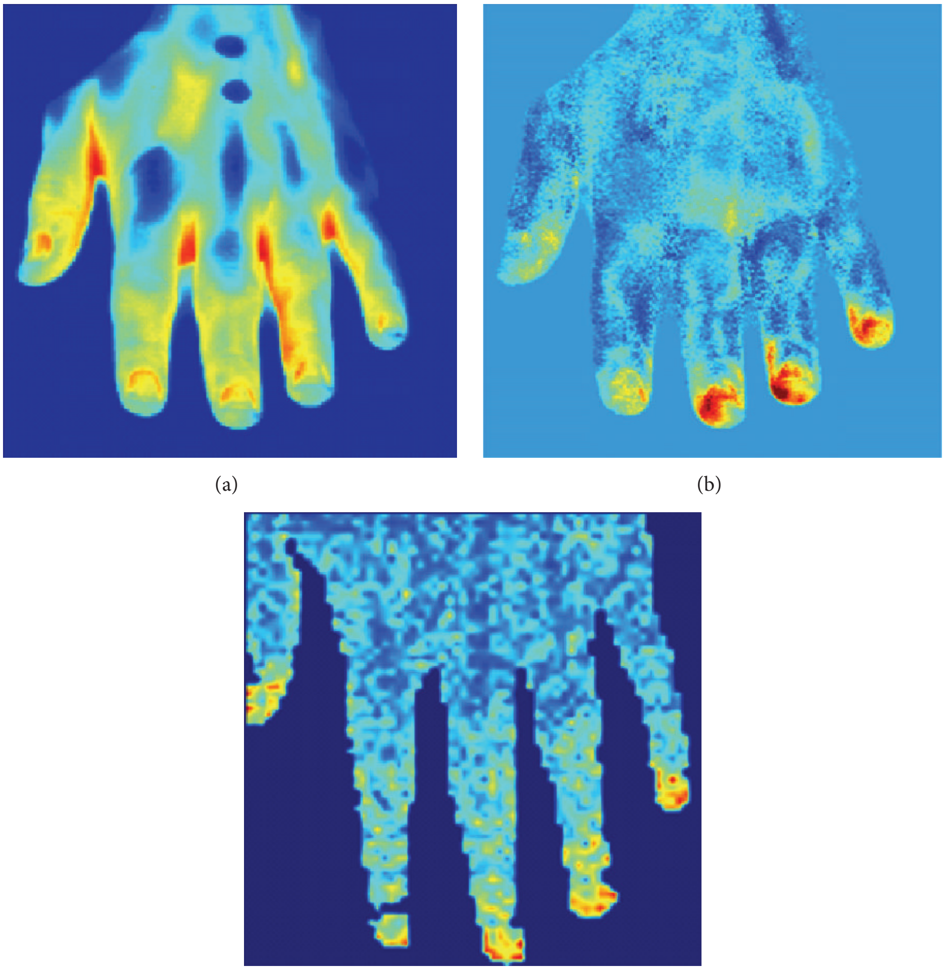
A number of studies suggest that the identification of active eccrine sweat glands by thermal imaging may have utility as a psychophysiological measure of sudomotor activity and may serve as a surrogate for the SCR when a contact method is either unavailable or undesirable [2, 6, 10, 12, 16, 36].

Recently, thermal IR imaging was used, together with standard GSR, to examine fear conditioning in posttraumatic stress disorder (PTSD) [37]. The authors examined fear processing in PTSD patients with mild symptoms and in individuals who did not develop symptoms, through the study of fear-conditioned response. The authors found that the analysis of facial thermal response during the conditioning paradigm performs like GSR to detect sympathetic responses associated with the disease.

2.5. Stress Response. An almost exclusive feature of thermal IR imaging in stress research is its noninvasiveness. Focused on professional drivers, a study of occupational ergonomics assessed mental workload using thermal IR imaging. Participants were exposed to simulator driving tasks both in the city and on the highway while cognitively challenged with a mental loading task (MLT). Compared with temperatures of the predriving session (baseline), significant differences were observed in the nose temperature across all conditions. The MLT seemed to have a defining effect on the temperature decrease of the nose, during the simulated city drive. No significant changes were observed on the forehead [38].

In a recent study, Pavlidis and colleagues [13] tried to quantify stress by measuring transient perspiratory responses on the perinasal area through thermal imaging. These responses proved to be sympathetically driven and, hence, a likely indicator of stress processes in the brain. The authors were able to monitor stress responses in the context of surgical training.

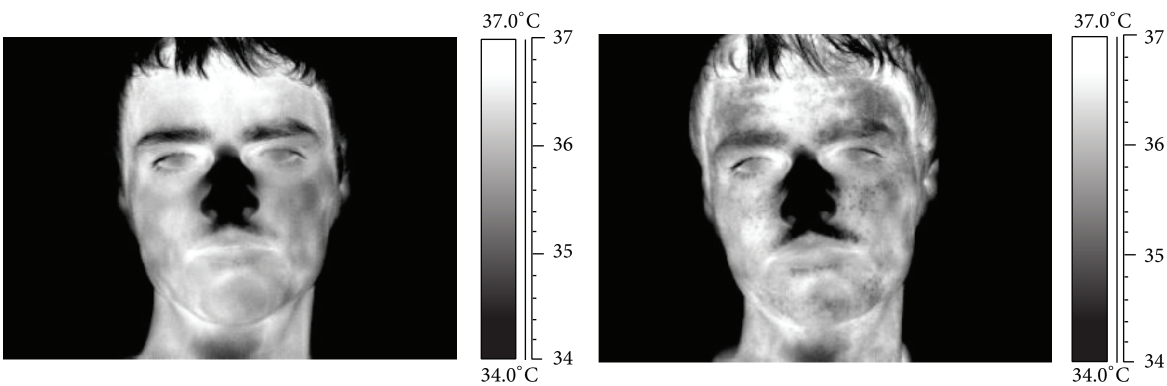
In another case and particularly in human-computer interaction field, Puri et al. [39] and Zhu et al. [40] used a Stroop task to exploit signs of frustration. Based on frontal regions, they observed that, compared with rest, stress increased blood volume into supraorbital vessels. Thermal IR imaging has also been used to assess affective training times by monitoring the cognitive load through facial temperature changes [41]. Learning proficiency patterns were based on an alphabet arithmetic task. Significant correlations, ranging



(a) (b)

(c)

Figure 4: Computation of cutaneous blood perfusion from thermal image series. (a) Thermal image of healthy hand; (b) cutaneous perfusion computed from thermal imagery (in arbitrary units); (c) laser Doppler image (in arbitrary units). The overall distributions appear to be consistent, both images similarly showing the same high-perfusion and low-perfusion regions.



(a) (b)

Figure 5: Emotional sweating and sudomotor response. The delivery of emotional pressure or stress stimulation (b) changes the rest of the (a) temperature distribution. The spotted dark signature is associated with the activity of the sweating glands.

from −0.88 to 0.96, were found between the nose tip temperature and the response time, accuracy, and the Modified Cooper Harper Scale ratings. Thermal IR thus represents a sensitive tool to assess learning and workload.

Engert et al. [15] explored the reliability of thermal IR imaging in the classical setting of human stress research. Thermal imprints were compared to established stress markers (heart rate, heart rate variability, finger temperature, alpha-amylase, and cortisol) in healthy subjects participating in two standard and well-established laboratory stress tests: the cold pressor test [42] and the trier social stress test [43]. Both tests showed evidence of thermal responses of several regions of the face. Although the thermal imprints and established stress marker outcome correlated weakly, the thermal responses correlated with stress-induced mood changes. On the contrary, the established stress markers did not correlate with stress-induced mood changes. These results suggest that thermal IR imaging provides an effective technique for the estimation of sympathetic activity in the field of stress research.

3. Discussion

Thermal IR imaging is a reliable method for ubiquitous and automatized monitoring of physiological activity. It provides a powerful and ecological tool for computational physiology. The reliability and validity of this method were proven by comparing data simultaneously recorded by thermal imaging and by golden standard methods, as piezoelectric pulse meter for pulse monitoring, piezoelectric thorax stripe for breathing monitoring or nasal thermistors, skin conductance, or galvanic skin response (GSR). As for the latter, studies have demonstrated that fIRI and GSR have a similar detection power [12, 13, 15, 35, 37]. Such results rely on the impressive advancement of the technology for thermal IR imaging. Modern devices ensure a high spatial resolution (up to 1280 × 1024 pixels with up to a few milliradiants in the field-of-view), high temporal resolution (full-frame frequency rate up to 150 Hz), and high thermal sensitivity (up to 15 mK at 30°C) in the spectral range $3 \div 5\,\mu$m [44]. The commercial availability of 640 × 480 focal plane array of uncooled and stabilized sensors (spectral range $7.5 \div 13.0\,\mu$m; full-frame frequency rate around 30 Hz; thermal sensitivity around 40 mK at 30°C) permits integrating this technology into automated systems for remote and automatic monitoring of physiological activity.

Real-time processing of thermal IR imaging data and data classification for psychophysiological applications is possible as the computational demand is not larger than that required for 640 × 480 pixels visible-band imaging data [2, 18, 44].

Thermal IR imaging has been indicated as a powerful tool to create, given the use of proper classification algorithms, an atlas of the thermal expression of psychophysiological responses [45, 46]. This would be based on the characterization of the thermal signal in facial regions of autonomic valence (nose or nose tip, perioral or maxillary areas, periorbital and supraorbital areas associated with the activity of the periocular and corrugator muscle and forehead), to monitor the modulation of the autonomic activity. Several studies have already shown the possibility of using thermal IR imaging in psychophysiology (see [2, 47] for reviews). These studies cover a number of fields, including developmental psychology and maternal empathy [48–50], social psychology [15, 51], and up to lie detection [52, 53].

However, several limitations exist for using thermal IR imaging in a real world. Because of the homeostasis, the cutaneous temperature is continuously adjusted to take into account the environmental conditions. Countermeasures must therefore be adopted to avoid attributing any psychological valence to pure thermoregulatory or acclimatization processes [2].

Also, despite the advantages offered by thermal IR imaging, it has to be taken into account that thermal signal development as a result of vascular change, perspiration, or muscular activity is rather slow with respect to other established techniques. Proper considerations should therefore be taken when monitoring thermal expression of psychophysiological activity.

Despite these limits, there is the concrete possibility of monitoring, in a realistic environment, at a distance and, unobtrusively, several physiological parameters and affective states. This opens the way for remote monitoring of the physiological state of individuals without requiring their collaboration and without interfering with their usual activities, thus suggesting the possibility of adding information of psychophysiological valence to behavioral or other typologies of investigation. One still unexplored but intriguing aspect is the study of possible correlation between individual thermal signatures and psychometric indexes, in order to assess, for example, whether given personality traits lead to interindividual differences in the facial thermal signature of autonomic activity or affective state or whether specific thermal expressions of specific personality or sociality traits exist. Of course, thermal IR imaging is not the first and unique attempt to explore these possibilities [54, 55], but thermal IR imaging seems to be one of the most ecological ones in this perspective. As such, thermal IR imaging provides an extraordinary opportunity to add physiological information to psychometric features, toward more robust classification of the individual's affective states, emotional responses, and profile.

A major issue that needs to be addressed for the practical application of thermal IR imaging in support of psychometrics concerns the adequacy of the method for identifying specific emotional or affective state at individual level. There are no specific studies available at the moment to answer this relevant question, which needs to be addressed by further research. A global limitation is derived from the fact that cutaneous thermal activity is intimately linked to the autonomic activity. The question therefore turns into "how specific and peculiar of each emotion are the autonomic responses and their thermal expression?" A definitive answer to this question is currently not available.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] R. Kirby, J. Forlizzi, and R. Simmons, "Affective social robots," *Robotics and Autonomous Systems*, vol. 58, no. 3, pp. 322–332, 2010.

[2] A. Merla, "Thermal expression of intersubjectivity offers new possibilities to human-machine and technologically mediated interactions," *Frontiers in Psychology*, vol. 5, article 802, 2014.

[3] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *Science*, vol. 221, no. 4616, pp. 1208–1210, 1983.

[4] R. Vetrugno, R. Liguori, P. Cortelli, and P. Montagna, "Sympathetic skin response," *Clinical Autonomic Research*, vol. 13, no. 4, pp. 256–270, 2003.

[5] S. D. Kreibig, "Autonomic nervous system activity in emotion: a review," *Biological Psychology*, vol. 84, no. 3, pp. 394–421, 2010.

[6] A. Merla, L. Di Donato, P. M. Rossini, and G. L. Romani, "Emotion detection through functional infrared imaging: preliminary results," *Biomedizinische Technick*, vol. 48, pp. 284–286, 2004.

[7] I. Fujimasa, "Pathophysiological expression and analysis of far infrared thermal images," *IEEE Engineering in Medicine and Biology Magazine*, vol. 17, no. 4, pp. 34–42, 1998.

[8] R. Murthy and I. Pavlidis, "Noncontact measurement of breathing function," *IEEE Engineering in Medicine and Biology Magazine*, vol. 25, no. 3, pp. 57–67, 2006.

[9] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 8, pp. 1418–1426, 2007.

[10] A. Merla and G. L. Romani, "Thermal signatures of emotional arousal: a functional infrared imaging study," in *Proceedings of the 29th Annual International Conference of IEEE-EMBS, Engineering in Medicine and Biology Society (EMBC '07)*, pp. 247–249, Lyon, France, August 2007.

[11] I. T. Pavlidis, J. Dowdall, N. Sun, C. Puri, J. Fei, and M. Garbey, "Interacting with human physiology," *Computer Vision and Image Understanding*, vol. 108, no. 1-2, pp. 150–170, 2007.

[12] D. Shastri, A. Merla, P. Tsiamyrtzis, and I. Pavlidis, "Imaging facial signs of neurophysiological responses," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 2, pp. 477–484, 2009.

[13] I. Pavlidis, P. Tsiamyrtzis, D. Shastri et al., "Fast by nature—how stress patterns define human experience and performance in dexterous tasks," *Scientific Reports*, vol. 2, article 305, 2012.

[14] A. Merla, "Method and system for the control of the residual efficiency of the interaction man-vehicle," European Patent EP13425145, 2013.

[15] V. Engert, A. Merla, J. A. Grant, D. Cardone, A. Tusche, and T. Singer, "Exploring the use of thermal infrared imaging in human stress research," *PLoS ONE*, vol. 9, no. 3, Article ID e90782, 2014.

[16] A. T. Krzywicki, G. G. Berntson, and B. L. O'Kane, "A noncontact technique for measuring eccrine sweat gland activity using passive thermal imaging," *Inernational Journal of Psychophysiology*, vol. 94, no. 1, pp. 25–34, 2014.

[17] C. B. Cross, J. A. Skipper, and D. Petkie, "Thermal imaging to detect physiological indicators of stress in humans," in *Thermosense: Thermal Infrared Applications XXXV*, vol. 8705 of *Proceedings of SPIE*, Baltimore, Md, USA, 2013.

[18] J. Dowdall, I. T. Pavlidis, and P. Tsiamyrtzis, "Coalitional tracking in facial infrared imaging and beyond," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2006.

[19] Y. Zhou, P. Tsiamyrtzis, and I. Pavlidis, "Tissue tracking in thermo-physiological imagery through spatio-temporal smoothing," *Medical Imaging Computing and Computer Assisted Intervention*, vol. 12, part 2, pp. 1092–1099, 2009.

[20] N. Sun, I. Pavlidis, M. Garbey, and J. Fei, "Harvesting the thermal cardiac pulse signal," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006*, vol. 4191 of *Lecture Notes in Computer Science*, pp. 569–576, Springer, Berlin, Germany, 2006.

[21] T. Bourlai, P. Buddharaju, I. Pavlidis, and B. Bass, "On enhancing cardiac pulse measurements through thermal imaging," in *Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine (ITAB '09)*, Larnaca, Cyprus, November 2009.

[22] A. A. Farag and S. Y. Chekmenev, "U.S. Patent Application 13/720,453," 2012.

[23] S. Y. Chekmenev, A. A. Farag, and E. A. Essock, "Thermal imaging of the superficial temporal artery: an arterial pulse recovery model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–6, IEEE, June 2007.

[24] A. Kamal, "Assessment of autonomic function in patients with rheumatoid arthritis using spectral analysis and approximate entropy method," *Neurosciences*, vol. 12, no. 2, pp. 136–139, 2007.

[25] J. N. Murthy, J. van Jaarsveld, J. Fei et al., "Thermal infrared imaging: a novel method to monitor airflow during polysomnography," *Sleep*, vol. 32, no. 11, pp. 1521–1527, 2009.

[26] G. F. Lewis, R. G. Gatto, and S. W. Porges, "A novel method for extracting respiration rate and relative tidal volume from infrared thermography," *Psychophysiology*, vol. 48, no. 7, pp. 877–887, 2011.

[27] R. Murthy, I. Pavlidis, and P. Tsiamyrtzis, "Touchless monitoring of breathing function," in *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '04)*, vol. 1, pp. 1196–1199, IEEE, September 2004.

[28] S. Y. Chekmenev, H. Rara, and A. A. Farag, "Non-contact, wavelet-based measurement of vital signs using thermal imaging," in *Proceedings of the 1st International Conference on Graphics, Vision, and Image Processing (GVIP '05)*, pp. 107–112, Cairo, Egypt, 2005.

[29] J. Fei and I. Pavlidis, "Thermistor at a distance: unobtrusive measurement of breathing," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 988–998, 2010.

[30] L. J. Goldman, "Nasal airflow and thoracoabdominal motion in children using infrared thermographic video processing," *Pediatric Pulmonology*, vol. 47, no. 5, pp. 476–486, 2012.

[31] J. Fei, I. Pavlidis, and J. Murthy, "Thermal vision for sleep apnea monitoring," *Medical Imaging Computing and Computer Assisted Intervention*, vol. 12, no. 2, pp. 1084–1091, 2009.

[32] A. Merla, L. di Donato, G. L. Romani, M. Proietti, and F. Salsano, "Comparison of thermal infrared and laser doppler imaging in the assessment of cutaneous tissue perfusion in healthy controls and scleroderma patients," *International Journal of Immunopathology and Pharmacology*, vol. 21, no. 3, pp. 679–686, 2008.

[33] I. Pavlidis and J. A. Levine, "Thermal image analysis for polygraph testing," *IEEE Engineering in Medicine and Biology Magazine*, vol. 21, no. 6, pp. 56–64, 2002.

[34] I. Fujimasa, T. Chinzei, and I. Saito, "Converting far infrared image information to other physiological data: the correlation of skin-surface temperature to physiological functions that control body temperature," *IEEE Engineering in Medicine and Biology*, vol. 19, no. 3, pp. 71–76, 2000.

[35] M. Coli, L. Fontanella, L. Ippoliti, and A. Merla, "Multiresolution KLE of psycho-physiological signals," in *Proceedings of S.Co. 2007, Book of Short Papers*, pp. 116–121, 2007.

[36] A. Merla, "Computational physiology in a thermal image setting," in *Proceedings of 5th Conference on Complex Models and Computational Intensive Methods for Estimation and Prediction (S.Co. '07)*, Book of Short Papers, pp. 338–343, Venice, Italy, 2007.

[37] A. Di Giacinto, M. Brunetti, G. Sepede, A. Ferretti, and A. Merla, "Thermal signature of fear conditioning in mild post traumatic stress disorder," *Neuroscience*, vol. 266, pp. 216–223, 2014.

[38] K. I. Calvin and V. G. Duffy, "Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement," *Occupational Ergonomics*, vol. 7, pp. 83–94, 2007.

[39] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren, "Stresscam: non-contact measurement of users' emotional states through thermal imaging," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, vol. 2, pp. 1725–1728, April 2005.

[40] Z. Zhu, P. Tsiamyrtzis, and I. Pavlidis, "Forehead thermal signature extraction in lie detection," in *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '07)*, pp. 243–246, Lyon, France, August 2007.

[41] J. Kang, J. A. McGinley, G. McFadyen, and K. Babski-Reeves, "Determining learning level and effective training times," in *Proceedings of the 25th Army Science Conference*, Orlando, Fla, USA, Novmber 2006.

[42] J. Hines and G. E. Brown, "A standard stimulant for measuring vasomotor reactions: its application in the study of hypertension," *Proceedings of the Staff Meetings of the Mayo Clinic*, vol. 7, pp. 332–335, 1932.

[43] C. Kirschbaum, K.-M. Pirke, and D. H. Hellhammer, "The "Trier social stress test"—a tool for investigating psychobiological stress responses in a laboratory setting," *Neuropsychobiology*, vol. 28, no. 1-2, pp. 76–81, 1993.

[44] P. Buddharaju, I. T. Pavlidis, and P. Tsiamyrtzis, "Physiology-based face recognition," in *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS '05)*, pp. 354–359, September 2005.

[45] M. M. Khan, R. D. Ward, and M. Ingleby, "Classifying pretended and evoked facial expressions of positive and negative affective states using infrared measurement of skin temperature," *ACM Transactions on Applied Perception*, vol. 6, no. 1, article 6, 2009.

[46] B. R. Nhan and T. Chau, "Classifying affective states using thermal infrared imaging of the human face," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 979–987, 2010.

[47] S. Ioannou, V. Gallese, and A. Merla, "Thermal infrared imaging in psychophysiology: potentialities and limits," *Psychophysiology*, vol. 51, no. 10, pp. 951–963, 2014.

[48] S. J. Ebisch, T. Aureli, D. Bafunno, D. Cardone, G. L. Romani, and A. Merla, "Mother and child in synchrony: thermal facial imprints of autonomic contagion," *Biological Psychology*, vol. 89, no. 1, pp. 123–129, 2012.

[49] B. Manini, D. Cardone, S. J. H. Ebisch, D. Bafunno, T. Aureli, and A. Merla, "Mom feels what her child feels:thermal signatures of vicarious autonomic response while watching children in a stressful situation," *Frontiers in Human Neuroscience*, 2013.

[50] S. Ioannou, S. Ebisch, T. Aureli et al., "The autonomic signature of guilt in children: a thermal infrared imaging study," *PLoS ONE*, vol. 8, no. 11, Article ID e79440, 2013.

[51] C. A. Hahn, R. D. Whitehead, M. Albrecht, C. E. Lefevre, and D. I. Perret, "Hot or not? Thermal reactions to social contact," *Biology Letters*, vol. 8, no. 5, pp. 864–867, 2012.

[52] I. Pavlidis and J. Levine, "Thermal image analysis for polygraph testing," *IEEE Engineering in Medicine and Biology Magazine*, vol. 21, no. 6, pp. 56–64, 2002.

[53] D. A. Pollina, A. B. Dollins, S. M. Senter et al., "Facial skin surface temperature changes during a 'concealed information' test," *Annals of Biomedical Engineering*, vol. 34, no. 7, pp. 1182–1189, 2006.

[54] K. Narita, T. Murata, T. Hamada et al., "Interactions among higher trait anxiety, sympathetic activity, and endothelial function in the elderly," *Journal of Psychiatric Research*, vol. 41, no. 5, pp. 418–427, 2007.

[55] M. Mendolia, "An index of self-regulation of emotion and the study of repression in social contexts that threaten or do not threaten self-concept," *Emotion*, vol. 2, no. 3, pp. 215–232, 2002.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 985032, 10 pages
http://dx.doi.org/10.1155/2015/985032

Research Article

Computational Psychometrics in Communication and Implications in Decision Making

Pietro Cipresso,[1] Daniela Villani,[2] Claudia Repetto,[2] Lucia Bosone,[3] Anna Balgera,[4] Maurizio Mauri,[4] Marco Villamira,[4] Alessandro Antonietti,[2] and Giuseppe Riva[1,2]

[1] *Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Via Ariosto 13, 20145 Milan, Italy*
[2] *Department of Psychology, Catholic University of the Sacred Heart, Largo Gemelli 1, 20123 Milan, Italy*
[3] *Social Psychology Research Group (EA 4163), Université Lumière Lyon 2, 5 avenue Pierre Mendès-France, 69656 Bron, France*
[4] *IULM University-Milan, Via Carlo Bo 8, 20143 Milan, Italy*

Correspondence should be addressed to Pietro Cipresso; p.cipresso@auxologico.it

Received 12 December 2014; Accepted 21 June 2015

Academic Editor: Maria N. D. S. Cordeiro

Recent investigations emphasized the role of communication features on behavioral trust and reciprocity in economic decision making but no studies have been focused on the effect of communication on affective states in such a context. Thanks to advanced methods of computational psychometrics, in this study, affective states were deeply examined using simultaneous and synchronized recordings of gazes and psychophysiological signals in 28 female students during an investment game. Results showed that participants experienced different affective states according to the type of communication (personal versus impersonal). In particular, participants involved in personal communication felt more relaxed than participants involved in impersonal communication. Moreover, personal communication influenced reciprocity and participants' perceptions about trust and reciprocity. Findings were interpreted in the light of the Arousal/Valence Model and self-disclosure process.

1. Introduction

Recently attention has been focused on the study of the effects of different types of communication on behavioral trust and reciprocity. Typically, the trust game is one of the preferred contexts to study trust and reciprocity in decision-making situations [1, 2].

As far as the role of communication in decision making is concerned, it is possible to identify two different perspectives. On the one hand, Bicchieri and colleagues defined *communication effect* [3, 4] as the positive effect of face-to-face (FtF) communication on cooperation and, particularly, on prosocial behavior. Specifically, oral communication led to the *identification* (humanization) of other agents. When another subject is perceived as being similar to ourselves—even if the similarity is vague and generic—we have the tendency to be kinder and more generous than we would be if the other part were completely anonymous. Identification reduces what is called "social distance" and therefore may increase the scope for reputation effects, which in turn may yield more cooperation [5]. In particular, Bicchieri [3, 6] distinguished between "relevant" and "irrelevant" communication. The former refers to the strategic discussion of the game and promise-making, usually producing more trust and cooperation than "irrelevant" communication [7].

Within the game-irrelevant communication, Buchan and colleagues [8] introduced a distinction between "personal" (e.g., subjects introduced themselves and talked about birthdays) and "impersonal" (e.g., subjects answered questions from the world almanac). Authors found that personal, game-irrelevant communication has a powerful influence on trusting behavior and they explained this result by stating that "the mere act of communicating more about themselves on a personal topic prompted participants to be significantly more concerned with others" [8].

As far as we know, the effects of communication on decision have been studied only in terms of behavioral results, such as the amount of money sent by the proposer (trust) and the amount of money sent by the responder (reciprocity). No studies have been focused on what happens during the communication process between participants in terms of affective and cognitive states.

To investigate this process we used computational psychometrics that has several advantages. First, this approach allows extracting several critical pieces of information related to subjective processes, without asking directly individuals. Second, thanks to the availability of brand-new technologies and paradigms researchers can investigate communication without interfering with the communication flow. Third, the use of computational psychometrics allows overcoming limits related to the use of the classic psychometrics questionnaires, where individuals are asked to recall *a posteriori* the communication experience.

2. Psychophysiological Correlates of Communication

At our best knowledge no study investigated yet the relationship between affective states and communication type. On the other hand several studies investigated the relationship between affective states and psychophysiological correlates in individuals by using biosensors [9].

To classify participant's affective states during communication, we referred to the well-known Arousal/Valence Model [10, 11], based on two dimensions: physiological arousal (high versus low) and emotional valence (positive versus negative). Lang and colleagues [12] found close relations between specific psychophysiological measures (e.g., corrugator and zygomatic electromyographic activity) and the self-reported valence dimension of emotion and between other psychophysiological measures (i.e., electrodermal activity) and the self-reported arousal dimension.

This approach has been extensively used in psychophysiological research as an objective way to measure affective states during a mediated experience [13–23]. More recently an extensive research has been done to discern different emotions by means of cardiovascular measures [24] and this result hugely helps the analysis of affective states in terms of cardiovascular indexes patterns [25–27]. For example, in a recent study Causse and colleagues [28] examined pilots' cardiovascular and oculometric measurements related to decision making.

To identify the affective and cognitive states with regard to the type of communication, we followed the process tracing paradigm [29] and we used an integrated psychophysiological approach. This approach is based on two methods frequently used to study behaviors and emotions, such as psychophysiological correlates [30] and eye movements recording [31–34]. Thanks to this integrated psychophysiological approach we aimed to overcome two principal limits related to the use of self-reports and behavioral data for assessment. At first, we would potentially index important processes which underlie affective responsivity but are not accessible to consciousness [35]. Secondly, we were interested in assessing emotional responding during the communication without interfering with the process. In this sense, psychophysiological measures allow us to continuously measure the temporal course of affective responding. The more common alternative of obtaining self-reports immediately after the episode is completed does not assess the temporal course of affective responsivity and may not even be particularly accurate indices of individuals' integrated responsivity over the course of the episode [36, 37].

3. Goal and Hypotheses

The general goal of this study was to investigate the effects of personal and impersonal communication on participants' affective states during the communication phase and their implications for trust, reciprocity, and participants' perception about trust and reciprocity after the trust game. For this reason we used an investment game as a typical decision-making task and we considered only the affective states during the communication phase (i.e., the period in which participants discuss options before making the decision). This is the first study carried out by using double eye-tracker synchronized with double psychophysiological biosensors.

In particular, we formulate two hypotheses:

(H1) There are significant differences between participants' affective states depending on the topic of communication (personal versus impersonal); since no study investigated this specific issue we decided to explore this relationship by referring to the Lang Model.

(H2) There are significant differences between participants' behaviors in terms of trust and reciprocity depending on the topic of communication (personal versus impersonal). According to Buchan et al. [8], we expected higher level of trust and reciprocity in personal communication condition. Furthermore, we would like to investigate also the perceptions related to the choice after having made the decision and to do that we analyzed participants' indications through their answers to specific questions.

4. Materials and Methods

4.1. Ethics Statement. Ethical Committee of the Psychology Department of the Catholic University of Milan approved the study. All participants gave written informed consent to the experimental procedure according to the rules of the scientific review board.

4.2. Research Design. The experimental investigation relies on the investment game, introduced by Berg and colleague [38], which provides a nice environment for the observation of trusting behavior in the lab. The between-subject design was based on the communication content factor suggested by Buchan et al. [8] (personal communication, PC, versus impersonal communication, IC). In the PC condition participants were allowed to discuss matters regarding personal interests, identity, ideas, and so forth. In the IC condition participants were limited to discuss arguments not linked to personal issues, such as the weather, public transportation, and life in the city where they lived.

4.3. Participants. Thirty-two students (mean age = 23.21 yrs.; SD = 2.38; range = 21–33) attending the Faculty of Psychology at the Catholic University of Sacred Heart of Milan, Italy, took part in the experiment. All the participants were students in psychology or communication, but without previous knowledge of decision science and in particular with no knowledge of the investment game used in the experiment. Social, economic, cultural, and historical background were all similar since all the participants were students in the same university and with similar background. To avoid differences in money perception the level of incomes was similar (student workers have been excluded). They were first met by one of the researchers during academic courses and then contacted via mail and/or telephone to schedule a meeting at the psychophysiology laboratory. The topics associated with the experiment were not mentioned during academic courses.

We selected only women in order to reduce the effects due to confounding variables associated with gender, such as sexual attraction, and control a priori effect of identification [39].

The participants were informed that the reimbursement for their attendance was related to their choices in the game.

Four of the participants failed to complete the experimental session due to personal or technical problems and were excluded from statistical analyses. The final sample was composed of 28 students, who were randomly coupled and assigned to one of the two experimental conditions (PC versus IC).

4.4. Procedure. The experimenters attached the following biosensors: a respiration belt (RSP), positioned on the chest; a facial EMG on corrugator supercilii muscle (i.e., EMG-CS), placed on the forehead after the researcher cleaned the skin with abrasive paste; two EEG sensors located on the frontal lobes, on the right and on the left, respectively, attached with conductive paste to improve signal detection (standard position on FP1 and FP2 of the EEG 10–20 international electrode system), with the reference electrodes attached to the respective ear lobes after they had been cleaned with abrasive paste by the experimenter; two SC adhesive patches applied on the nondominant palm; a BVP sensor placed on top of the index finger of the nondominant hand.

Finally, participants' eye movements were calibrated by asking participants to follow a small point on the screen. The EEG and the physiological signals of BVP, GSR, respiration, and EMG-CS were acquired using a ProComp Infiniti device from Thought Technology, including Biograph Infiniti 5.0.2 software to record all signals. Every channel was synchronously acquired at 2048 Hz and exported at a 256 Hz sampling rate (every 3.90625 milliseconds).

In order to analyze deeper the affective states related to different types of communication content during the communication phase in FtF condition, we used video-conferencing technologies. According to Brosig et al. [39], video-conferencing is useful for employing the favorable features of FtF communication as a "real" conference and it produces cooperation rates very close to FtF communication [40].

Both members of each pair were separately welcomed in different rooms by two experienced experimenters, who assisted them for the duration of the laboratory session. The experimenters were instructed to maintain a neutral voice tone and a neutral behavior while the subjects were being exposed to experimental stimuli. Participants did not know each other and could not talk to each other before the beginning of the experiment. First of all they were informed about the general purpose of the research (summarized in the sentence: "The goal of the experiment is to analyze the different ways people take a decision") and received all the instructions about the nature of the investment game (including the switch of the roles). Then, they signed the written informed consent and were prepared in order to collect the psychophysiological data. When a participant indicated feeling comfortable, the researcher asked her to try to remain still during all the experimental conditions in order to avoid artifacts in signal acquisitions as a result of movements. A baseline measure of the psychophysiological parameters was obtained with a 5-min registration in a steady state [27].

At this point the experimental task started. The two participants were allowed to get to know each other, meeting their partner in the virtual context of Google video chat. The communication lasted 3 min. They were asked to converse about specific arguments among authorized topics, according to the experimental condition to which they belonged (PC versus IC).

After this step, the decision phase started and participants knew that the game was played with real monetary payoffs and the winnings were converted in a proportioned way.

The proposer was asked to choose how much money she would give to the partner from an initial amount of 2000 Euros by choosing one of the following alternatives: 100, 500, 1000, 1500, and 2000 Euros. According to the investment game procedure, the responder was then asked to choose how much of the triple of the amount received she wanted to give back to the proposer, selecting one of the following options: nothing, 1/4, 1/2, or 3/4 of the final amount and all the final amount.

After the decision, the video chat was closed and participants completed a brief questionnaire about the choice they had made and related personal motivations and the partner's choice and related motivations.

The second phase of the experiment started with another 3-min video-chat communication, when the game was repeated, by inverting the roles.

4.5. Signal Acquisition and Data Analysis. The experiment was carried out in two labs, each one equipped with two portable PCs, one for delivering the stimuli and acquiring eye-tracker data and the other for psychophysiological signal recording.

The physiological signals were acquired using a ProComp Infiniti device from Thought Technology, including Biograph Infiniti 5.0.2 software to record and export all raw signals. Every channel was synchronously acquired at 2048 Hz and exported at least a 256 Hz sampling rate (every 3.90625 milliseconds) or higher where needed.

The pupillometry data were acquired using two Tobii x-series, including Tobii Studio software to record all raw signals, then exported, and resampled at 60 Hz.

In our study, by using the eye-tracker data extraction, we obtained for each participant a matrix of gaze and pupil data corresponding to stimuli presentation (such as the webpage for Google video chat); in particular, we collected 50 rows for each second (sampling to 50 Hz), thereby making it possible to establish the exact period previously indicated.

Then, the synchronization of the psychophysiological signals allowed us to identify the communication period. In this case, we used several algorithms to synchronize eye-tracker systems with a ProComp Infiniti device from Thought Technology by using the TT-AV Sync sensor, which was configured through a physical channel on ProComp capturing the light (e.g., by identifying black and white) thanks to a photodiode actually applied on the screen [41]. Moreover, based on gazes and pupil signals acquired, it has been possible to identify eye-blinks, which enabled us to align the matrix containing the eye-blink data from gazes and pupil signals, with the matrix containing the psychophysiological signals. Thanks to this procedure, we correctly identified the communication period with an error of ±0.01 seconds.

All collected biosignals were analyzed using Matlab 7.0 (The Mathworks, Natick, MA, USA) for the signal processing and computation of psychophysiological measures.

Data were analyzed with the aid of the statistical software SPSS, version 18 (Statistical Package for the Social Sciences—SPSS for Windows, Chicago, IL, USA).

4.6. Psychophysiology and Affective States. According to the classic valence-arousal model [10, 11] described in Introduction, we considered the two dimensions of physiological arousal and emotional valence for identifying affective states in participants during the experimental session.

Physiological arousal can be measured by using electroencephalogram (EEG), galvanic skin response (GSR), cardiovascular activity (ECG or BVP), and respiration signal (RSP); emotional valence can be measured by using EEG, self-reports, facial expression identification, eye-blink startle, and facial EMG corrugator and/or zygomatic. According to Blumenthal and colleagues [42], facial EMG-CS (corrugator) can be considered the best measure for emotion valence.

Frontal EEG activation asymmetry has been generally used, giving evidences that greater left frontal activity seems to be higher related to positive emotional valence, whereas greater right frontal activity seems to be more involved in negative emotional valence [43]. Alpha index seems to be the most adaptation to study the frontal EEG activation asymmetry [43]. Thus, Alpha Asymmetry index can be calculated in many different ways to take into account the hemispheric prevalence and to correct the sign accordingly. In calculating this index it is crucial to consider that higher cortical activation is revealed by lower Alpha waves, and thus this needs to be taken into account in the computation and formula derivation.

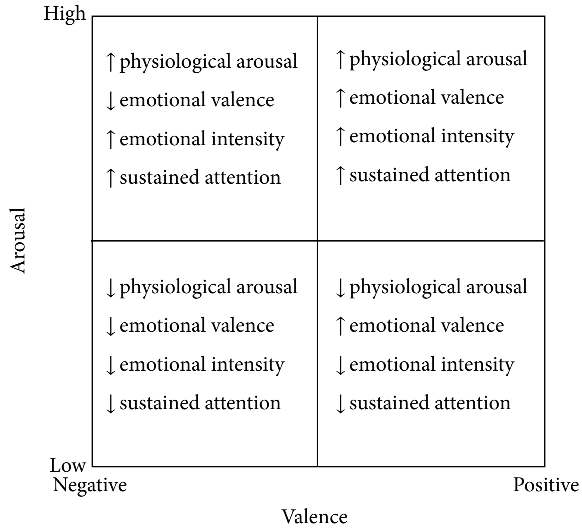


Figure 1: The affective space, defined by the dimensions valence and arousal, based on Lang [10] and the physiological activation in central or peripheral nervous system.

Furthermore, we used pupillometry to investigate emotional responses. Because pupil dilation indicates a higher cognitive load or emotional processing [25, 27], studies have suggested that combining pupillometry and the use of functional magnetic resonance imaging (fMRI) to produce a pupil dilation index can provide summative information on cognitive and affective processes [44]. Figure 1 provides the relationships between psychological affective states and physiological "activation" increasing (↑) and decreasing (↓).

As recently showed by Mojzisch and colleagues [45], the self-involvement during social interactions is specifically related to attention allocation. For this reason we are interested also in investigating the sustained attention as another dimension of analysis. Slow Alpha EEG bands (following Slow Alpha) from EEG have been demonstrated to be a valid measure of sustained attention [46, 47].

4.7. Signal Processing. Cardiovascular and respiratory activities were monitored to evaluate both voluntary and autonomic effect of respiration on heart rate, analyzing IBI (Interbeat-Interval extracted from Blood Volume Pulse sensor, recognized as a measure equivalent to R-R interval extracted from electrocardiogram), respiration (from chest strip sensor), and their interaction. Following the guidelines of task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, typical Heart Rate Variability (HRV) temporal and spectral method indexes were used to evaluate the autonomic nervous system response [48–52]. Temporal indexes were computed from IBI as indicated in the guideline as heart rate (HR), number of pairs of successive NNs (derived from IBIs) that differ by more than 50 ms (millisecond), namely, NN50, and Root Mean Square Standard Deviation (RMSSD). Spectral analysis was performed using Fourier spectral methods with custom software. The rhythms have been classified as very

Table 1: Description of variables and their measures (↑ is for increase; ↓ is for decrease).

Variable	Measures
↑physiological arousal	↑heart rate/↑skin conductance/↑EEG Beta waves
↑emotional valence	↓EMG corrugator/↓EEG Alpha Asymmetry
↑emotional intensity	↑pupil dilation
↑sustained attention	↓EEG Slow Alpha waves
↑negative affect and activation	↑LF/HF/↓NN50/↓RMSSD

low frequency (VLF, < 0.04 Hz), low frequency (LF, from 0.04 to 0.15 Hz), and high frequency (HF, from 0.15 to 0.5 Hz) oscillations. This procedure enabled us to calculate the LF/HF ratio, also known as the sympathovagal balance index. More, we were able to check if variations in ECG signal were due to variation in respiration affected by talking. An accurate checking of LF, respiration rate, and a simple analysis of respiratory sinus arrhythmia allowed us to estimate no effects of the talking on ECG signals. All artifacts in ECG signals have been corrected, according to the guidelines [52] and Clifford and colleagues [53].

SC and skin resistance recorded through GSR sensors are units of electrodermal activity, which are expressed in either conductance (microsiemens) or resistance (microohms). SC reflects a fairly slow physiological process and can be sampled at 32 Hz without distortion. The signal is expressed in microohms. We considered mean and standard deviation of the sampled signal.

The raw electromyography is a collection of positive and negative electrical signals. Their frequency and amplitude gave us information on the contraction or rest state of the muscle. Amplitude was measured in μV (microvolts). As the participant contracts the muscle, the number and amplitude of the lines increase; as the muscle relaxes, it decreases. We considered the Root Mean Square (RMS) for rectifying the raw signal of the corrugator supercilii muscle and converting it to an amplitude envelope, following EMG-CS. We were not interested in frequency related to this muscle fatigue.

EEG signals need to be extensively worked to remove ocular artifacts and blinks using automatic algorithm and subsequent visual inspection. Then the corrected matrixes have been computed to calculate means of the Slow Alpha bands (at 7–10 Hz) for the FP1 and FP2 EEG channel, through spectral analyses [54–56]. Eye-tracker usefulness in this study was also critical for synchronizing the psychophysiological signals within the communication phase (based on a photodiode applied on the eye-tracker monitor and connected as a channel to the psychophysiology control unit). Table 1 offers a clear indication of the variables measured as described above.

4.8. Questions Asked after the Decision. After the decision, the video chat was closed and participants completed a brief questionnaire. In particular, proposers were asked about their choice and the personal motivations for it through two questions:

(i) How do you consider the amount of money you gave to the responder? ("Big" or "Small");

(ii) Why did you give to your partner this amount of money? ("Because she seemed trustworthy," "Because she did not seem trustworthy," "Because she seemed generous," or "Because she did not seem generous").

Responders too were asked about their choice and the personal motivations for it through two questions:

(i) How do you consider the amount of money you gave back to the proposer? ("Fair," "Partially fair," or "Unfair");

(ii) Why did you give to your partner this amount of money? ("Because it was the right choice," "Because she would have done the same," "Because the gain had to be shared equally between me and her," "Because at the beginning it was her money," and "Because thanks to me the money have been increased").

5. Results

The dependent variables considered were defined as follows:

(i) Psychophysiological measures (extracted following the above indication; see 2.7): heart rate, EMG corrugator, skin conductance, EEG Alpha Asymmetry, EEG Slow Alpha waves, EEG Beta waves, LF/HF, pupil dilation, NN50, and RMSSD.

(ii) Trust: the amount of money sent by the proposer to the responder.

(iii) Reciprocity: the amount of money returned by the responder to the proposer.

(iv) Perceived trust: the individual's evaluation of fairness of personal decision about the amount of money sent to the responder.

(v) Perceived reciprocity: the individual's evaluation of fairness of personal decision about the amount of money sent to the proposer.

Analyses were carried out in order to test the hypotheses. For this reason we analyzed differences between participants' affective states depending on the topic of communication (H1) and differences between participants' trust and reciprocity and their perceptions about trust and reciprocity depending on the topic of communication (H2).

5.1. Comparison between Participants' Affective States Depending on the Topic of Communication (H1). We analyzed the effect played by the independent variable "role" (proposer versus responder) on all dependent variables: since it resulted not statistically significant in both main and interaction effects, it was not further considered in the subsequent analyses.

The impact of communication type on psychophysiological measures of affective states was evaluated by running a series of independent samples t-tests in a between-factor design (PC versus IC) to compare the two groups

Table 2: Descriptive statistics of physiological measures for the two conversation groups (PC versus IC).

Measure	Condition	N	Mean	Std. deviation	Std. error
Heart rate*	Personal	14	1.169537	0.282227	0.075428
(*physiological arousal*)	Impersonal	14	1.492253	0.514155	0.137414
EMG corrugator*	Personal	14	1.308825	0.296176	0.079156
(*emotional valence*)	Impersonal	14	1.704306	0.590298	0.157764
Skin conductance*	Personal	14	1.310103	0.305998	0.081782
(*physiological arousal*)	Impersonal	14	1.744871	0.644443	0.172235
EEG Alpha Asymmetry*	Personal	14	1.619759	0.331382	0.088566
(*emotional valence*)	Impersonal	14	2.034656	0.613816	0.164049
EEG Slow Alpha waves*	Personal	14	1.655140	0.281301	0.075181
(*sustained attention*)	Impersonal	14	1.996572	0.529739	0.141579
EEG Beta waves*	Personal	14	1.270383	0.331340	0.088554
(*physiological arousal*)	Impersonal	14	1.656981	0.619189	0.165485
LF/HF	Personal	14	1.230078	0.272284	0.072771
(*Heart Rate Variability*)	Impersonal	14	1.474582	0.429780	0.114864
Pupil dilation*	Personal	14	.4252	.09369	.02504
(*emotional intensity*)	Impersonal	14	.5309	.16404	.04384
NN50	Personal	14	0.699334	0.132247	0.035344
(*Heart Rate Variability*)	Impersonal	14	0.603127	0.173188	0.046286
RMSSD*	Personal	14	0.674883	0.306916	0.082027
(*Heart Rate Variability*)	Impersonal	14	0.372111	0.422183	0.112833

*$P \leq .055$ according to corrected t-tests showed in Table 3.

Table 3: Physiological measures: between condition effects (corrected t-tests).

	Communication type (personal versus impersonal communication)			
	Mean difference	t_{26}	St. Err.	Sig.
Heart rate	−0.3227	−2.059	0.1567	.053
EMG corrugator	−0.3955	−2.241	0.1765	.037
Skin conductance	−0.4348	−2.280	0.1907	.035
EEG Alpha Asymmetry	−0.4149	−2.225	0.1864	.038
EEG Slow Alpha waves	−0.3414	−2.130	0.1603	.046
EEG Beta waves	−0.3866	−2.060	0.1877	.053
LF/HF	−0.2445	−1.798	0.1360	.084
Pupil dilation	−0.2031	−2.029	0.1001	.046
NN50	0.0962	−1.652	0.0582	.111
RMSSD	0.3028	−2.170	0.1395	.039

of participants. t-test correction has been done whether or not dependent variable variance between groups was homogeneous, based on Levene's test.

The descriptive statistics summarized in Table 2 and the corrected t-tests in Table 3 showed that participants in the PC condition had lower HR and SC, indicating a lower physiological arousal than participants in the IC condition. On the other hand lower EMG corrugator and lower Alpha Asymmetry indexes for participants in the PC condition indicated a positive emotional valence but a negative emotional valence in participants in the IC condition. Moreover, results showed a lower sympathovagal balance index (LF/HF ratio) and higher RMSSD and NN50. This last triplet of indexes, with these trends, still demonstrated a state characterized by positive valence and low arousal in the PC condition and a state characterized by negative valence and high arousal in IC condition. Pupil dilation showed statistical significant differences between conditions, indicating that emotional intensity was higher in IC. Negative emotions had a stronger effect on states characterized by negative valence and high arousal than positive emotions on states characterized by positive valence and low arousal did.

Table 4: Money sent by the proposer to the responder.

Valid	Personal communication		Impersonal communication	
	Frequency	Percent	Frequency	Percent
100	0	0	2	14.3
500	1	7.1	3	21.4
1000	7	50	3	21.4
1500	4	28.6	2	14.3
2000	2	14.3	4	28.6
Total	14	100	14	100

Table 5: Money returned by the responder to the proposer.

Valid	Personal communication		Impersonal communication	
	Frequency	Percent	Frequency	Percent
1/4	1	7.1	3	21.4
1/2	6	42.9	9	64.3
3/4	7	50	2	14.3
Total	14	100	14	100

Finally, Slow Alphas were lower in participants in the PC condition, highlighting their higher sustained attention than in IC.

A power analysis was used to anticipate the likelihood that our study yielded significant effects. The goal of our post hoc power analysis was to compute achieved power, given significance level, effect size, and sample size. The analysis showed a good power (above 80% per each measure). These results also demonstrated that our sample is at the fair size.

5.2. Comparison between Participants' Trust and Reciprocity Depending on the Topic of Communication and Their Perceptions about Trust and Reciprocity after the Choice (H2). The impact of communication type on trust and reciprocity was evaluated by running a series of independent t-tests (see Tables 4 and 5 for frequencies). No differences were found between the amount of money sent from the proposer to the responder in the PC and IC conditions (t = 0.577, df = 26, and P = .569). As far as reciprocity was concerned, results suggested that responders returned a higher amount of money to the proposers in the PC condition (t = 2.096, df = 26, and P = .046). Nobody gave back zero or the total final amount.

We investigated also the participants' perceptions about their choices by using Kendall tau-c and we compared them depending on the communication condition.

Analyses showed that differences exist between participants' perceptions about trust depending on the topic of communication. In particular, the motivations related to the proposer's play were different between PC and IC (tau_c = −0.384, df = 26, and P = .032): the proposers in PC condition considered the responders more trustworthy than in IC condition.

As a consequence, responders evaluated the personal decision about the amount of money returned to the proposers as fairer in the PC condition (tau_c = 0.459, df = 26, and P = .020).

6. Discussion and Conclusions

According to Bicchieri [3, 6], the content of communication plays a key role during the interaction in decision making. In particular, according to Buchan et al. [8], we decided to concentrate our study on the "irrelevant communication" by differentiating between "personal" irrelevant communication condition and "impersonal" irrelevant communication condition.

The general aim of this study was to investigate the effects of personal and impersonal communication on participants' affective states during the communication phase preceding a decision and on their perception about trust and reciprocity after the investment game. A cutting-edge innovation was represented by the computational psychometrics approach, using double eye-tracker synchronized with double psychophysiological biosensors.

As far as the first hypothesis is concerned, we found differences between participants' affective states depending on the topic of communication. We also verified that psychophysiological correlates of affective states experienced during the communication phase remained stable and were confirmed during the decision phase.

In particular, according to the Arousal/Valence Model [10, 11], psychophysiological measures showed two different profiles. Participants in the PC condition were more relaxed than participants in the IC condition and positive emotional valence, low arousal, and high attentional resources characterized their affective states. On the other hand, participants in the IC condition were more activated than participants in the other condition and low emotional valence, high arousal, and low attentional resources characterized their affective states.

Probably participants who were allowed to speak about personal interests, identity, and ideas felt low activated and experienced positive states. It is possible to place these individuals in the affective dimension of positive valence and low arousal of the Lang Model (see Figure 1). On the other hand, participants who were limited to discussing arguments not linked to personal issues felt highly activated and uncomfortable. It is possible to place these persons in the affective dimension of negative valence and high arousal of the Lang Model (see Figure 1).

This sense of being comfortable could be explained by the fact that in personal communication condition participants were involved in a self-disclosure process. Self-disclosure has been defined as any message about the self that an individual communicates to another [57, 58]. According to traditional interpersonal theories, self-disclosure is a type of communication through which individuals make themselves known to other people and when others reciprocate by sharing revealing information it leads to intimacy and relational development [59]. Furthermore, according to the uncertainty reduction theory, individuals will not only seek information to reduce uncertainty but also reciprocate with similar amounts of information and at the same level of intimacy [60]. The self-disclosure process could have got participants in personal communication condition closer to each other and more comfortable.

Psychophysiological correlates supported the self-disclosure process also regarding sustained attention. Indeed, participants in the PC condition showed high sustained attention even if their other states fitted with the relax dimension. Participants belonging to the impersonal communication showed low sustained attention even if their other states fitted with the negative valence and high arousal affective state. Probably the personal content of communication led participant to feel engaged in friendly and attentive interaction with the partner [61].

According to our results, the second hypothesis was partially supported. The feeling of being relaxed and more comfortable during personal communication was not associated with proposer's trust, even if proposers in PC considered the responders as trustworthy. However, the positive emotional valence, low arousal, and high attentional resources featured by participants in PC condition were associated with a higher amount of money returned to the proposer and an evaluation about the amount of money returned to the proposers as fairer than in the IC condition.

To summarize, we found differences between psychophysiological correlates of communication conditions related to the affective and cognitive states. Furthermore, we found a significant difference related to the choices and participants' perceived trust and reciprocity according to the content of communication.

Many different issues are still open and deserve further investigation. First of all, although we selected only women in order to reduce the effects due to confounding variables associated with gender, future researches should include men in order to verify whether gender differences exist. Second, since the investigation of the relationships between the affective states during personal versus impersonal communication and the behavioral responses was out of the purpose of the present research, we did not try to pinpoint causal links between the two. However, the next step would be to build predictive models able to identify the affective states that lead to differences in trusting behavior. Third, we opted for a single-shot trust game instead of iterative one. This choice fitted our experimental goals and a future challenge would be to repeat the experimental setting in multiple rounds of the trust games in order to investigate the effects of repeated communications on trust: does the deepening of the self-disclosure (in PC) versus the repeated talk about generic topics (in IC) change the trust over time?

The present results suggest that a personal conversation is more oriented to positive affective states and to positive perceptions related to the partner of communication. Implications could be relevant for many purposes, since many communication processes made in both FtF and online settings can precede decision making. Vivid examples are related to the physician-patient communication [62, 63], both FtF and online, where the physician can be rated as more empathic thanks to the open communication used [64] or also the insurance and financial fields, where online decision making is rapidly becoming more relevant than ever. In both cases, a personal communication could help patients and clients, such as the responders of real-life decisional situations, to feel more comfortable and this affective state could lead them to be more favorably disposed during the decisional process.

Thanks to the use of computational psychometrics we have been able to investigate communication processes and their effects on decision making. In particular, the use of double systems synchronized to record the related psychophysiological correlates and eye movements has shown great potential.

The extended use of new computational psychometrics paradigms represents a promising approach and future researches aiming to investigate communication processes are encouraged to follow these advances methods.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] A. Tversky and D. Kahneman, "Judgment under uncertainty: heuristics and biases," *Science*, vol. 185, no. 4157, pp. 1124–1131, 1974.

[2] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA, 1944.

[3] C. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Cambridge University Press, New York, NY, USA, 2006.

[4] C. Bicchieri and A. Lev-On, "Computer-mediated communication and cooperation in social dilemmas: an experimental analysis," *Politics, Philosophy & Economics*, vol. 6, no. 2, p. 139, 2007.

[5] M. M. Krasnow, L. Cosmides, E. J. Pedersen, and J. Tooby, "What Are Punishment and Reputation for?" *PLoS ONE*, vol. 7, no. 9, Article ID e45662, 2012.

[6] C. Bicchieri, "Covenants without swords: group identity, norms, and communication in social dilemmas," *Rationality and Society*, vol. 14, no. 2, pp. 192–228, 2002.

[7] K. S. Bouas and S. S. Komorita, "Group discussion and cooperation in social dilemmas," *Personality and Social Psychology Bulletin*, vol. 22, no. 11, pp. 1144–1150, 1996.

[8] N. R. Buchan, E. J. Johnson, and R. T. A. Croson, "Let's get personal: an international examination of the influence of communication, culture and social distance on other regarding preferences," *Journal of Economic Behavior and Organization*, vol. 60, no. 3, pp. 373–398, 2006.

[9] B. Colombo, C. Rodella, S. Riva, and A. Antonietti, "The effects of lies on economic decision making. An eye-tracking study," *Research in Psychology and Behavioral Sciences*, vol. 1, no. 3, pp. 38–47, 2013.

[10] P. J. Lang, "The emotion probe. Studies of motivation and attention," *American Psychologist*, vol. 50, no. 5, pp. 372–385, 1995.

[11] J. A. Russell, "Affective space is bipolar," *Journal of Personality and Social Psychology*, vol. 37, no. 3, pp. 345–356, 1979.

[12] P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm, "Looking at pictures: affective, facial, visceral, and behavioral reactions," *Psychophysiology*, vol. 30, no. 3, pp. 261–273, 1993.

[13] R. W. Backs, S. P. da Silva, and K. Han, "A comparison of younger and older adults' self-assessment Manikin ratings of affective pictures," *Experimental Aging Research*, vol. 31, no. 4, pp. 421–440, 2005.

[14] M. M. Bradley and P. J. Lang, "Affective reactions to acoustic stimuli," *Psychophysiology*, vol. 37, no. 2, pp. 204–215, 2000.

[15] J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, *Handbook of Psychophysiology*, Cambridge University Press, Cambridge, UK, 3rd edition, 2007.

[16] P. Cipresso, S. Serino, D. Villani et al., "Is your phone so smart to affect your state? An exploratory study based on psychophysiological measures," *Neurocomputing*, vol. 84, pp. 23–30, 2012.

[17] D. Grühn and S. Scheibe, "Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (LAPS): do ratings become more extreme with age?" *Behavior Research Methods*, vol. 40, no. 2, pp. 512–521, 2008.

[18] A. Keil, M. M. Bradley, O. Hauk, B. Rockstroh, T. Elbert, and P. J. Lang, "Large-scale neural correlates of affective picture processing," *Psychophysiology*, vol. 39, no. 5, pp. 641–649, 2002.

[19] B. Kuhr, J. Jacobi, C. Krause et al., "Arousal, valence, dominance... and desire? Evidence from an Erp study concerning the necessity of a new motivational dimension to describe affective states," *Psychophysiology*, vol. 48, p. S88, 2011.

[20] J. D. Morris, "Observations: SAM: the self-assessment manikin—an efficient cross-cultural measurement of emotional response," *Journal of Advertising Research*, vol. 35, no. 6, pp. 63–68, 1995.

[21] B. Rozenkrants and J. Polich, "Affective ERP processing in a visual oddball task: arousal, valence, and gender," *Clinical Neurophysiology*, vol. 119, no. 10, pp. 2260–2265, 2008.

[22] R. B. Rubin, A. M. Rubin, E. E. Graham, E. M. Perse, and D. R. Seibold, "Self-assessment Manikin," in *Communication Research Measures II: A Sourcebook*, pp. 336–341, 2009.

[23] C. F. Sharpley, P. Kamen, M. Galatsis, R. Heppel, C. Veivers, and K. Claus, "An examination of the relationship between resting heart rate variability and heart rate reactivity to a mental arithmetic stressor," *Applied Psychophysiology Biofeedback*, vol. 25, no. 3, pp. 143–153, 2000.

[24] P. Rainville, A. Bechara, N. Naqvi, and A. R. Damasio, "Basic emotions are associated with distinct patterns of cardiorespiratory activity," *International Journal of Psychophysiology*, vol. 61, no. 1, pp. 5–18, 2006.

[25] V. Magagnin, M. Mauri, P. Cipresso et al., "Heart rate variability and respiratory sinus arrhythmia assessment of affective states by bivariate autoregressive spectral analysis," in *Proceedings of the Computing in Cardiology (CinC '10)*, pp. 145–148, September 2010.

[26] M. Mauri, P. Cipresso, A. Balgera, M. Villamira, and G. Riva, "Why is Facebook so successful? Psychophysiological measures describe a core flow state while using Facebook," *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 12, pp. 723–731, 2011.

[27] M. Mauri, V. Magagnin, P. Cipresso et al., "Psychophysiological signals associated with affective states," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 3563–3566, IEEE, Buenos Aires, Argentina, September 2010.

[28] M. Causse, B. Baracat, J. Pastor, and F. Dehais, "Reward and uncertainty favor risky decision-making in pilots: evidence from cardiovascular and oculometric measurements," *Applied Psychophysiology Biofeedback*, vol. 36, no. 4, pp. 231–242, 2011.

[29] N. Reisen, U. Hoffrage, and F. W. Mast, "Identifying decision strategies in a consumer choice situation," *Judgment and Decision Making*, vol. 3, no. 8, pp. 641–658, 2008.

[30] A. Bechara and A. R. Damasio, "The somatic marker hypothesis: a neural theory of economic decision," *Games and Economic Behavior*, vol. 52, no. 2, pp. 336–372, 2005.

[31] A. T. Duchowski, "A breadth-first survey of eye-tracking applications," *Behavior Research Methods, Instruments, and Computers*, vol. 34, no. 4, pp. 455–470, 2002.

[32] G. L. Lohse and E. J. Johnson, "A comparison of two process tracing methods for choice tasks," *Organizational Behavior and Human Decision Processes*, vol. 68, no. 1, pp. 28–43, 1996.

[33] J. E. Russo and F. Leclerc, "An eye-fixation analysis of choice processes for consumer nondurables," *Journal of Consumer Research*, vol. 21, no. 2, pp. 274–290, 1994.

[34] J. E. Russo and L. D. Rosen, "An eye fixation analysis of multialternative choice," *Memory & Cognition*, vol. 3, no. 3, pp. 267–276, 1975.

[35] A. J. Tomarken, "A psychometric perspective on psychophysiological measures," *Psychological Assessment*, vol. 7, no. 3, pp. 387–395, 1995.

[36] J. T. Cacioppo, J. S. Martzke, R. E. Petty, and L. G. Tassinary, "Specific forms of facial EMG response index emotions during an interview: from Darwin to the continuous flow hypothesis of affect-laden information processing," *Journal of Personality and Social Psychology*, vol. 54, no. 4, pp. 592–604, 1988.

[37] B. L. Fredrickson and D. Kahneman, "Duration neglect in retrospective evaluations of affective episodes," *Journal of Personality and Social Psychology*, vol. 65, no. 1, pp. 45–55, 1993.

[38] J. Berg, J. Dickhaut, and K. McCabe, "Trust, reciprocity, and social history," *Games and Economic Behavior*, vol. 10, no. 1, pp. 122–142, 1995.

[39] J. Brosig, J. Weimann, and A. Ockenfels, "The effect of communication media on cooperation," *German Economic Review*, vol. 4, no. 2, pp. 217–241, 2003.

[40] N. Bos, D. Gergle, J. S. Olson, and G. M. Olson, "Being there versus seeing there: trust via video," in *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHI '01)*, pp. 291–292, Seattle, Wash, USA, April 2001.

[41] P. Cipresso, M. Mauri, A. Balgera, E. Roman, and M. A. Villamira, "Synchronization of a biofeedback system with an eye tracker through an audiovisual stimulus marker," *Applied Psychophysiology and Biofeedback*, no. 35, p. 332, 2010.

[42] T. D. Blumenthal, B. N. Cuthbert, D. L. Filion, S. Hackley, O. V. Lipp, and A. Van Boxtel, "Committee report: guidelines for human startle eyeblink electromyographic studies," *Psychophysiology*, vol. 42, no. 1, pp. 1–15, 2005.

[43] S. Debener, A. Beauducel, K. Fiehler, S. Rabe, and B. Brocke, "Frontal EEG alpha asymmetry and affective style: are individual differences related to fundamental dimensions of emotion?" *Psychophysiology*, vol. 38, p. S35, 2001.

[44] G. J. Siegle, S. R. Steinhauer, V. A. Stenger, R. Konecky, and C. S. Carter, "Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data," *NeuroImage*, vol. 20, no. 1, pp. 114–124, 2003.

[45] A. Mojzisch, L. Schilbach, J. R. Helmert, S. Pannasch, B. M. Velichkovsky, and K. Vogeley, "The effects of self-involvement on attention, arousal, and facial expression during social interaction with virtual others: a psychophysiological study," *Social neuroscience*, vol. 1, no. 3-4, pp. 184–195, 2006.

[46] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Research Reviews*, vol. 29, no. 2-3, pp. 169–195, 1999.

[47] W. Klimesch, M. Doppelmayr, H. Russegger, T. Pachinger, and J. Schwaiger, "Induced alpha band power changes in the human EEG and attention," *Neuroscience Letters*, vol. 244, no. 2, pp. 73–76, 1998.

[48] N. D. Ahuja, A. K. Agarwal, N. M. Mahajan, N. H. Mehta, and H. N. Kapadia, "GSR and HRV: its application in clinical diagnosis," in *Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems*, New York, NY, USA, June 2003.

[49] R. Barbieri, J. K. Triedman, and J. P. Saul, "Heart rate control and mechanical cardiopulmonary coupling to assess central volume: a systems analysis," *The American Journal of Physiology—Regulatory Integrative and Comparative Physiology*, vol. 283, no. 5, pp. R1210–R1220, 2002.

[50] R. M. Kramer, "Trust and distrust in organizations: emerging perspectives, enduring questions," *Annual Review of Psychology*, vol. 50, pp. 569–598, 1999.

[51] R. J. Lewicki, E. C. Tomlinson, and N. Gillespie, "Models of interpersonal trust development: theoretical approaches, empirical evidence, and future directions," *Journal of Management*, vol. 32, no. 6, pp. 991–1022, 2006.

[52] M. Malik, J. T. Bigger, A. J. Camm et al., "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[53] G. D. Clifford, F. Azuaje, and P. McSharry, "ECG statistics, noise, artifacts, and missing data," in *Advanced Methods and Tools for ECG Data Analysis*, chapter 3, pp. 55–99, Artech House, London, UK, 2006.

[54] A. I. Bagic, R. C. Knowlton, D. F. Rose, and J. S. Ebersole, "American Clinical Magnetoencephalography Society clinical practice guideline 1: recording and analysis of spontaneous cerebral activity," *Journal of Clinical Neurophysiology*, vol. 28, no. 4, pp. 348–354, 2011.

[55] A. I. Bagić, R. C. Knowlton, D. F. Rose, and J. S. Ebersole, "American clinical magnetoencephalography society clinical practice guideline 3: MEG-EEG reporting," *Journal of Clinical Neurophysiology*, vol. 28, no. 4, pp. 362–363, 2011.

[56] V. V. Nikulin and T. Brismar, "Long-range temporal correlations in alpha and beta oscillations: effect of arousal level and test-retest reliability," *Clinical Neurophysiology*, vol. 115, no. 8, pp. 1896–1908, 2004.

[57] L. R. Wheeless, "A follow-up study of the relationships among trust, disclosure, and interpersonal solidarity," *Human Communication Research*, vol. 4, no. 2, pp. 143–157, 1978.

[58] L. R. Wheeless and J. Grotz, "Conceptualization and measurement of reported self-disclosure," *Human Communication Research*, vol. 2, no. 4, pp. 338–346, 1976.

[59] D. A. Taylor and I. Altman, "Communication in interpersonal relationships: social penetration processes," in *Interpersonal Processes: New Directions in Communication Research*, pp. 257–277, Sage, 1987.

[60] C. R. Berger and R. J. Calabrese, "Some explorations in initial interaction and beyond: toward a developmental theory of interpersonal communication," *Human Communication Research*, vol. 1, no. 2, pp. 99–112, 2006.

[61] E. E. Graham, C. A. Barbato, and E. M. Perse, "The interpersonal communication motives model," *Communication Quarterly*, vol. 41, no. 2, pp. 172–186, 1993.

[62] S. Riva, M. Monti, P. Iannello, and A. Antonietti, "The representation of risk in routine medical experience: what actions for contemporary health policy?" *PLoS ONE*, vol. 7, no. 11, Article ID e48297, 2012.

[63] S. Riva, M. Monti, P. Iannello, G. Pravettoni, P. J. Schulz, and A. Antonietti, "A preliminary mixed-method investigation of trust and hidden signals in medical consultations," *PLoS ONE*, vol. 9, no. 3, Article ID e90941, 2014.

[64] J. Silvester, F. Patterson, A. Koczwara, and E. Ferguson, "'Trust me...': psychological and behavioral predictors of perceived physician empathy.," *Journal of Applied Psychology*, vol. 92, no. 2, pp. 519–527, 2007.