IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 29, 2019, accepted October 7, 2019, date of publication November 4, 2019, date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2951349

Eliciting Contact-Based and Contactless Gestures With Radar-Based Sensors

NATHAN MAGROFUOCO[1], JORGE-LUIS PÉREZ-MEDINA[2], PAOLO ROSELLI[3,4], JEAN VANDERDONCKT[1], (Senior Member, IEEE), AND SANTIAGO VILLARREAL[1]

[1]Louvain Research Institute in Management and Organizations (LRIM). Université Catholique de Louvain (UCL), 1348 Louvain-la-Neuve, Belgium
[2]Intelligent and Interactive Systems Laboratory (SI^2 Lab), Universidad de Las Américas (UDLA), Quito 170504, Ecuador
[3]Matematica Dipartimento, Università degli Studi di Roma, 00173 Rome, Italy
[4]Institut de recherche en mathématique et physique, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium

Corresponding author: Jorge-Luis Pérez-Medina (jorge.perez.medina@udla.edu.ec)

ABSTRACT Radar sensing technologies now offer new opportunities for gesturally interacting with a smart environment by capturing microgestures via a chip that is embedded in a wearable device, such as a smartwatch, a finger or a ring. Such microgestures are issued at a very small distance from the device, regardless of whether they are contact-based, such as on the skin, or contactless. As this category of microgestures remains largely unexplored, this paper reports the results of a gesture elicitation study that was conducted with twenty-five participants who expressed their preferred user-defined gestures for interacting with a radar-based sensor on nineteen referents that represented frequent Internet-of-things tasks. This study clustered the $25 \times 19 = 475$ initially elicited gestures into four categories of microgestures, namely, micro, motion, combined, and hybrid, and thirty-one classes of distinct gesture types and produced a consensus set of the nineteen most preferred microgestures. In a confirmatory study, twenty new participants selected gestures from this classification for thirty referents that represented tasks of various orders; they reached a high rate of agreement and did not identify any new gestures. This classification of radar-based gestures provides researchers and practitioners with a larger basis for exploring gestural interactions with radar-based sensors, such as for hand gesture recognition.

INDEX TERMS Contact-based gesture, contactless gesture, gesture classification, gesture elicitation study, gestural interaction, microgesture, radar sensing.

I. INTRODUCTION

The increasing presence of electronic devices in our daily lives extends their field of application, thereby driving researchers to investigate new interaction modalities for making our lives easier. Major progress has been made in research on interactions via whole body or hand gestures in recent years. Contactless user interfaces enable end users to view, to control, and to manipulate any digital content, such as an object, an item, or a scene, without physically touching the user interface or the device. These interfaces are explored in a wide range of application scenarios, from medical surgery for hygiene to car dashboard controllers for safety. Vision-based recognition has achieved a high level of accuracy in stationary frameworks but remains challenging in mobile contexts of use, which are demanding in terms of illumination conditions. Contactless interfaces now benefit from a new source of input, namely, **radar sensing** [1]–[4].

Radio frequency overcomes some issues, but low spatial resolution remains a barrier. Technical options have become available for making dynamic finger/hand recognition feasible; these options include ultrawide bandwidth (UWB) for digital millimeter-wave radars, which combine a compact physical form-factor with high accuracy at short ranges with high frequency [1], low frequency [5], or a micro-Doppler signatures [6].

Although these techniques have reached a new level of operationalization, our knowledge about microgestures, namely, gestures that are issued at a very small distance from the radar, is still in its infancy; only a few gestures have been studied. This paper investigates this issue by exploring user-defined microgestures based on radar sensing for Internet of Things (IoT) tasks by eliciting these microgestures from participants via the classical research method of a gesture elicitation study [7], which is augmented with additional

The associate editor coordinating the review of this manuscript and approving it for publication was Hadi Heidari.

variables. A second study with new participants, in which gestures are selected from the previously described gesture set for more diverse tasks, confirms this classification.

II. RELATED WORK

This section is divided into three subsections: a presentation of the most significant radar sensing technologies for gestural input, an introduction to gesture elicitation studies with a brief overview of their scope, and an overview of previous studies on microgestures and related work. A *microgesture* is defined here as a gesture that has small articulation variations [8], [9], e.g., small movements such as translations and rotations (*micro-motions*), low speed, low acceleration, low pressure, small size, and short time.

A. RADAR-SENSING TECHNOLOGIES

Several systems support contactless interaction by gesture recognition. LLAP [10] tracks and localizes the end user by ultrasound to enable contactless gesture inputs. LLAP utilizes speakers and microphones that are embedded in smartphones to play and record sound waves that are inaudible to humans. By measuring the phase of the sound signal that is reflected by the hands or fingers of the user, LLAP measures gesture movements with an accuracy of up to 3.5 mm.

Strata [11] consists of a smartphone that transmits known audio signals at inaudible frequencies, and the received signal that is reflected by a moving finger is analyzed to track the finger location. Strata selects the channel tap that corresponds to the finger movement and extracts the phase change of the selected tap to accurately estimate the distance change of the finger. Strata estimates the absolute distance of the finger, thereby realizing high accuracy and low latency.

GestureDrawer [12] introduces a novel one-handed interaction technique that enables users to place imaginary user interface controls in 3D empty space without receiving any kind of feedback (completely eye-free and device-free) and without fixed landmarks.

RadarCat [2] is a small, versatile radar-based system for material and object classification that enables new forms of everyday proximate interaction with digital devices. The system is trained to recognize and classify various types of materials and objects in real time.

The sense of agency (SoA) [13] refers to the subjective experience of voluntary control over actions in the external world. This SoA is tested on contactless systems using the intentional binding paradigm by comparing contactless systems with physical interactions. The haptic feedback enables the exploration of how modalities influence intentional binding.

Google ATAP (Advanced Technology and Projects) has launched **Soli** (Fig. 1a), which is a new sensing technology that uses miniature radar to detect contactless gesture interactions[1]. This radar operates according to the principles of reflection and detection of radio-frequency electromagnetic waves [14]. This radar was initially used for various tasks, e.g., tracking large objects, such as cars and planes, but engineers wanted to track micro motions from twitches of human hands. The Soli sensor is manufactured as a compact semiconductor device, which requires low energy [3]. This microsystem is used for ubiquitous gesture interaction for a large variety of applications, which include Internet of Things (IoT), virtual reality (VR) [15], object selection and manipulation [2], material recognition [16], and creative musical and audio-visual contexts [17]. Soli is the first millimeter-wave radar system that is designed end-to-end for ubiquitous and intuitive fine gesture interaction. Soli was officially released to be incorporated in the forthcoming Google Pixel 4 smartphone for gesture recognition[2]. Despite these advances, only five system-defined gestures are investigated [3], [16]: single-tapping to press a virtual button (Fig. 1b), rubbing a finger to turn a virtual dial (Fig. 1c), moving the index finger to the left/right to move a virtual slider (Fig. 1d), and horizontal/vertical swiping for navigation (Fig. 1e-f).

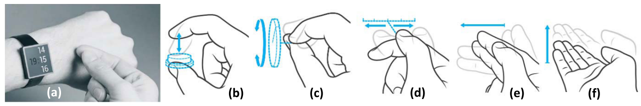


FIGURE 1. Google Soli sensor (a), coming with five system-defined gestures [15]: a virtual button (b), a virtual dial (c), a virtual slider (d), a horizontal (e) and a vertical (f) swipe.

Ens *et al.* [18] explore a multidimensional design space that combines microgestures that are captured by a Google Soli that is attached to a LeapMotion device with other types of gestures, which are detected by a belt sensor and a Microsoft HoloLens, within the greater lexicon of gestures. Via this approach, these authors expect to combine large-scale movements for gross-grained navigation with small-scale gestures for fine-grained interaction. Dynamic continuous hand gesture recognition that is based on a frequency-modulated continuous wave (FMCW) radar sensor has also been proven to be successful, especially because the radar system does not depend on lighting, noise, or atmospheric conditions [6], such as in Latern [19]. Dynamic hand gesture recognition is proved feasible with four dynamic gestures [5]: hand rotation, beckoning, snapping fingers, and flipping fingers.

Although the aforementioned technologies have become affordable for tracking microgestures in everyday objects, our knowledge of what kind of microgestures would be preferred and used by end users in various contexts of use remains limited.

B. GESTURE ELICITATION STUDIES

Understanding users' preferences and behaviors with a new interactive technology from the early stages of design empowers designers with valuable information for shaping a product's characteristics to realize more effective and efficient use. *Gesture Elicitation Studies* (GESs) [7], [20], [21]

[1] http://atap.google.com/soli/

[2] See https://9to5google.com/2019/06/11/google-pixel-4-project-soli/

are popular for investigating users' preferences for gesture input in a variety of contexts of use [7], [22]–[24].

The result consists of a characterization of users' gesture input behaviors with valuable information for designers, practitioners, and end users regarding the consensus among participants (which is computed as an *agreement* or *coagreement rates* [7], [20], [21]), most frequent (thus, generalizable across users) gestures for a specified task, and insights into users' conceptual models for performing tasks.

C. OVERVIEW OF GESTURE ELICITATION STUDIES

GESs are conducted along the three traditional dimensions of the *context of use* [25]: users and their tasks, devices and platforms they use, and physical environments [26].

1) ON VARIOUS DEVICES AND PLATFORMS

Since their inception, gesture elicitation studies have focused primarily on specified *platform or device*, such as tabletops in surface computing [7], [27], mobile interactions [22], and smart televisions [24], [28], where the entire device or a component, such as the trackpad, is considered [29]. In [28], the authors conduct user-elicited studies with hand gestures, in which the natural attitudes, preferences and memorability are considered, with the objective of generating and implementing a set of TV control commands. Chattopadhyay and Bolchini [30] propose the concept of motor-intuitive, which is a touchless interaction primitive (mid-air, directional strokes) that is based on the up-down and left-right space schemas for facilitating the support of intuitive touchless interactions and large-display touchless interactions. Gheran et al. [31] study small gestures that are issued while wearing a LogBar Zero ring device, some of which are mid-air gestures and others are microgestures with or without contact. However, despite these studies, radar-based sensors have not been subjected to any GESs.

2) IN DIFFERENT ENVIRONMENTS

Gestures are typically elicited in an *environment* in which devices are situated, such as the steering wheel in a car [32] or public displays, where social acceptance must be considered [33]. In these public displays, the tasks also vary; examples include selecting and buying a good in a store [34] and opt-in/out [33]. Space and place affect the engagement of participants [35].

3) FOR MULTIPLE USERS

A GES is said to be user-independent if it does not target any specified user profile or category. In contrast, a GES is *user-dependent* if a user profile, such as children who are issuing whole-body gestures, is devised. A GES can target any population of end users instead of a platform or environment. Different types of users may play different roles, which are also influenced by their culture [36]. [37] studied hand gestures, while [38] compared freehand gestures with gestures that are issued on the skin. Any limb, physical capability or deficiency thereof may also become the subject of a GES.

Only two studies are similar to our study. Chan *et al.* [9] introduced the notion of a single-hand microgesture (SHMG), which is issued with only one hand. That study involved 34 referents, which covered six categories: transform, simulation, browsing, editing, menu, and selection. This paper will investigate other categories of referents with only the simulation in common: volume up/down, mute, play, pause, and stop. Four categories emerge: tap, swipe, circle, and draw.

Bostan *et al.* [37] collected 957 microgestures from 19 participants for 26 referents that correspond to frequent smartphone tasks, such as select, navigate, open/close, minimize/maximize, zoom in/out, next/previous, switch task, scroll up/down, and accept/reject call. The results from that study demonstrate that end users tend to prefer one hand for holding an object, assign different meanings to different parts of the hand, and place higher importance on hand properties than on skin properties.

III. EXPERIMENT

A. PARTICIPANTS

Twenty-five voluntary participants (15 females and 10 males; aged 12 to 68 years; $M = 27.92$, $SD = 13.30$, and $Mdn = 23$) were recruited for the study via a contact list for various organizations. Twenty-three participants were right-handed, and 2 participants were left-handed. The occupations of the participants included secretary, teacher, director, psychologist, retired, and students in engineering, law, economics, physiotherapy, management, and criminology. All participants reported frequent use of computers and smartphones and no dexterity problems.

B. APPARATUS

The experiment for a radar-based sensor was conducted in a quiet office meeting room. A simple computer screen was used to display the referents to the participants. All the gestures were recorded by a camera that was placed in front of the participants to capture their hands and fingers without capturing their faces. To keep the study centered on the topic, the participants were asked to limit their movements to their hands and fingers without any other constraint.

C. PROCEDURE

1) PRE-TEST PHASE

The participants were welcomed to the setup by a researcher and were asked to sign a GDPR-compliant informed consent form. Then, the participants were given information about the study and the general process of the experiment. The participants were also asked to complete a sociodemographic questionnaire (e.g., age, gender, handedness, use of technologies based on a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree) and to perform a creativity test and a motor-skill test. The participants answered a series of questions on the "Test My Creativity"[3] and received an assessment of their level of creativity.

[3] http://www.testmycreativity.com/ is a web-based application that proposes a collection of hands-on tools and a systematic method that are designed to systematize the space of creativity, innovation and value creation.

Agreement rates that were computed by AGATe [20] for referents, which are sorted in decreasing order, with 95.

The motor-skill test [39] consisted of pinching each finger with the thumb several times. The questionnaire and the tests are intended to determine whether the participants satisfied minimal ability requirements for conducting the experiment and for future data analysis.

2) TEST PHASE

The participants watched a short video that explained the principle of radar-based sensing without showing any gestures; some participants were asked about gestures, and the tasks to perform were described to them. Each session implemented the original GES protocol [7]; participants were presented with 19 referents (Table 2, first column), namely, actions for controlling various objects in an IoT environment, for which the participants proposed suitable gestures for executing those referents, namely, gestures that fit the referents well. Microgestures are not frequent in gesture elicitation studies. Therefore, we believed that the legacy bias [40] would not substantially affect the gestures that were proposed by the participants. We applied *visual priming*, which is part of *framed guessability* [41], to provide the participants guidance regarding the scenario before and after the gesture should be proposed.

The Participants were instructed to remain as natural as possible. The global order of the referents was randomized per participant. When a member of a pair of related gestures, such as ''Turn TV On/Off'', was randomly selected, the other member came just afterwards so that participants could consider the relation.

The **Thinking time** between the first showing of the item and the moment when the participant knew which gesture she would perform was measured in seconds with a stopwatch. After performing each gesture, the researcher asked the participants to provide a **Goodness-of-fit** rating from 1 to 10 to express the extent to which they thought their gesture was suitable for the presented referent. Each session took approximately 45 minutes per participant. One researcher welcomed participants and facilitated the completion of the form, another presented the referents, and another supervised, thereby ensuring minimal interference from the researchers.

3) POST-TEST PHASE

Participants completed the IBM PSSUQ (post-study system usability questionnaire) [42], which enabled them to express their levels of satisfaction with the usability of this new device and the testing process. There is a high correlation ($\rho = .89$) between the results of this questionnaire and the perceived usability of any interactive system [43].

IV. RESULTS AND DISCUSSIONS

A. DESCRIPTION

Each elicited gesture is identified with its corresponding video, is subjected to *descriptive labeling* [44] based on Nielsen's procedure [45] and is assigned to a gesture category. An example of this process is presented in Table 2. We collected a total of 475 gesture proposals from 25 participants for 19 referents, from which 2 samples were removed as outliers because they were interrupted. We clustered the gestures into 4 groups and 2 subgroups of similar gestures according to the criteria that are presented below (as presented in Table 1). A classification was obtained according to the taxonomy that is presented in [30].

Microgesture: A microgesture falls into one of the two groups: (i) *bimanual*: bimanual microtap gestures (fingertips tapping on the surface of the hand, fingers and fingertips) and bimanual microslide gestures (fingertips sliding on the projected surface plane of the hand), and (ii) *finger rubs*: microgestures that involve at least two fingers rubbing against each other.

Motion gesture: A motion gesture is any gesture that uses a 3D object, pose, and motion to interact. This category includes any gesture change or motion. We focus mainly on spatial 3D motion gestures, which use a defined hand pose or property to define a cluster configuration. Hand poses define the state of the hand, palm, thumb, and individual fingers based on the *simple hand model* [24], which considers the 5 fingertips and the palm point.

Combination of gestures: Two gestures are combined to produce a new gestures.

Others: This category consists of recurring gestures that did not fit in any other categories but that were worth considering.

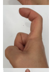
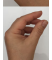
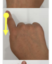
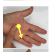
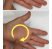
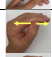
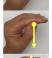
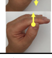
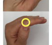
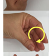
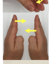
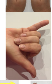
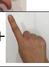
We also added another category of gestures (99): nonrecurring combinations of gestures that could not be linked to any existing category. In the end, we identified 31 distinct gesture categories. In the next section, we assign the gestures to meaningful categories and conduct a general analysis of the results, gesture types and consensus among gesture proposals.

B. CONSENSUS BETWEEN ELICITED GESTURES

Fig. 2 presents the agreement rates that are obtained for each referent in decreasing order, along with the corresponding most frequently elicited gestures. Overall, the agreement rates, as computed by AGATe [20], range from low to medium agreement, namely, from .05 to .19 ($M = .107$, $SD = .042$). The three highest agreement rates corresponded to the following actions: ''turn the light on'', (ref. 10) with an agreement rate of 19.30%; ''turn the light off'' (ref. 11), with an agreement rate of 17.70%; and ''go to the previous item in a list'' (ref. 7), with an agreement rate of 13%. The three lowest agreement rates corresponded to ''turn AC off'' (ref. 9), with an agreement rate of 4.3%; and to the pair ''turn alarm on/off'' (ref. 16 and 17), which both had an agreement rate of 5%. These were considered the least familiar and least frequent tasks.

These results are very similar to the lowest-ever-reported agreement rates in the literature on gesture elicitation ([20], p. 1332), according to a summary of the agreement

TABLE 1. Classification of gestures for radar-based sensors.

Group	Subgroup	Gesture	Gesture category	Description	Classification
Motion gesture	Spatial 3D		1: Uni-pinch	Index thumb pinch open; Index thumb pinch open; Index middle pinch open	Temporal-Static-Simple
			2: Multipinch	Uni-pinch repeated several times	Temporal-Static-Generalized
			3: Point	Index point; Middle point; Middle ring point	Temporal-Dynamic-Simple
			4: Thumb	Thumb up	Temporal-Static-Simple
			5: Hook	Index hook	Temporal-Static-Simple
			6: Flat	Flat push; Flat wave in; Flat wave out	Temporal-Static-Simple
			7: Drag	Point drag; Pinch drag; Trigger drag	Temporal-Dynamic-Generalized
			8: Rotate	Point rotate; Pinch rotate	Spatial-Non-Referential
			9: Scale	2 Point scale; 2 Pinch scale; 2 Trigger scale	Temporal-Dynamic-Generalized
			10: Swipe	Point swipe; Splay swipe	Temporal-Dynamic-Generalized
			11: Scroll	Point; Scroll	Temporal-Dynamic-Generalized
			12: Flick	Splay flick; Flat hand; Flick	Temporal-Dynamic-Generalized
			13: Tilt	Splay tilt; Flat hand; Tilt	Spatial-Referential
			14: Fist	Fist push; Fist hold	Spatial-Referential
			15: Splay	Splay push; Splay wave in/out	Spatial-Referential
			16: Wave	Wave hand	Spatial-Referential
Microgesture	Bimanual		17: Back tap	Black flat tap	Spatial-Referential
			18: Palm tap	Palm tap	Spatial-Referential
			19: Horizontal swipe	Palm horizontal swipe	Spatial-Referential
			20: Vertical swipe	Palm vertical swipe	Spatial-Referential
			21: Palm draw	Palm stroke circle	Temporal-Dynamic-Generalized
			22: Back slide	Back horizontal slide; Back vertical slide	Temporal-Dynamic-Generalized
	Finger rubs		23: Horizontal rub	Finger lateral rub	Temporal-Dynamic-Generalized
			24: Vertical rub	Finger backward rub	Temporal-Dynamic-Generalized
			25: Sagittal rub	Finger frontward rub	Temporal-Dynamic-Generalized
Group	Subgroup	Gesture	Gesture category	Description	Classification
Others	None		26: Snap	Finger snap	Temporal-Dynamic-Generalized
			27: Button	Thumb push on the index side	Temporal-Static-Generalized
			28: Dimmer	Index and thumb rotational movement	Spatial-Referential
			29: Clap	Hand clap	Spatial-Referential
			30: Phone	Thumb and auricular tense	Temporal-Dynamic-Generalized
Combination of gestures	None	 + 	31: Swipe + Point	Point swipe; splay swipe + index point; index middle point; index middle ring point	Spatial-Referential

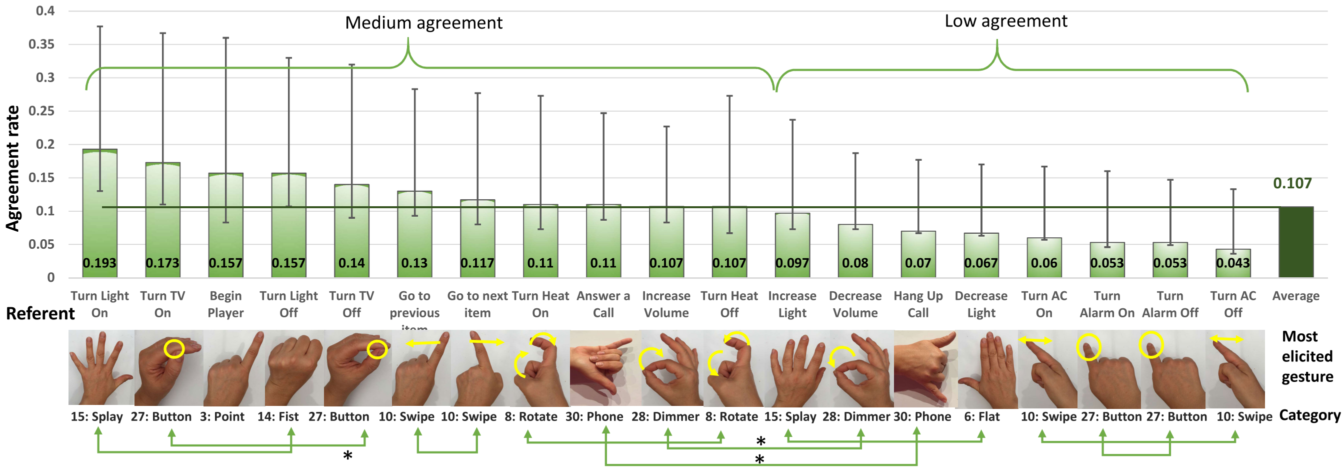


FIGURE 2. Agreement rates that were computed by AGATe [20] for referents, which are sorted in decreasing order, with 95% confidence intervals ($\alpha = .05$).

rates of 18 studies, for which the smallest value (.108) was reached by Seyed *et al.* [44] for multidisplay gestures. These low values are due to two factors: (i) the design space of possible gestures with a sensor such as Google Soli is large and contains many possibilities in $0D$, $1D$, $2D$, and $3D$; hence, participants tried many gestures; and (ii) *legacy bias* [27] is not strongly present in this case due to the recency and the novelty of the system. AGATe [20] also enables us to conduct a statistical test to determine whether their rates are correlated. Since our sampling contains several pairs of related referents, we computed the rates; however, only three pairs were significant: "Turn TV On/Off" ($V_{rd}(1, n = 50) = 5.556$, $p^{*} \leq .05$), "Increase/decrease volume" ($V_{rd}(1, n = 50) = 4.000$, $p^{*} \leq .05$), "Go to next/previous item" ($V_{rd}(1, n = 50) = .667$, $p > .05$, *n.s.*), "Turn AC On/Off" ($V_{rd}(1, n = 50) = 1.190$, $p > .05$, *n.s.*), "Turn light On/Off" ($V_{rd}(1, n = 50) = 2.283$, $p > .05$, *n.s.*), "Increase/decrease light" ($V_{rd}(1, n = 50) = 2.455$, $p > .05$, *n.s.*), "Turn heat On/Off" ($V_{rd}(1, n = 50) = .053$, $p > .05$, *n.s.*), "Turn alarm On/Off" ($V_{rd}(1, n = 50) = 0.0$, $p > .05$, *n.s.*), "Answer/hang up call" ($V_{rd}(1, n = 50) = 6.0$, $p^{*} \leq .05$). Only a few pairs had correlated gestures.

C. DESIGN SPACE OF RADAR-BASED MICROGESTURES

To better evaluate our participants' gesture proposals, six dimensions of analysis were considered, which were informed by the hand and finger gestures (Fig. 4):

- *Dimensionality:* This dimension expresses the dimension of a gesture and ranges from 1D (a line), 2D (a plane), 2D1/2 (a spatial 3D gesture that is projected onto a 2D surface) to 3D (in space). Surprisingly, 69% of the elicited gestures were issued in space and subsequently projected (20%), thereby suggesting that participants fully exploited the space around the device even if it was limited.
- *Range of motion:* This dimension relates the distance between the position of the human body that is producing the gesture and the location of the gesture. In our classification, gestures have been regarded as small for microgestures that only involve fingers, as medium if the distance is moderate (e.g., performed using a smartphone or a similar device), and as large if the gesture requires full-arm movements. Participants issued medium gestures in most cases (73%), followed by small (24 %), and large gestures (3%). Large gestures were almost nonexistent, since they are contrary to the nature of microinteractions. Medium gestures had a range of approximately 1-3 cm; therefore, they could be called centi-gestures, since they are on the order of centimeters. Small gestures were below this threshold, with a range of 5-9 mm; therefore, they could be called milligestures. These distances are critical for two reasons: the gesture location is required for differentiating the locus of action, e.g., the closest pinch gesture is related to a "On" command, while the farthest pinch gesture triggers a corresponding "Off" command; and the same gesture that is performed close to or far from the source will be perceived differently, as the level of the signal weakens as the target is moved away from the radar.
- *Body part:* A gesture was categorized differently if it was performed with the whole hand versus the fingers. For instance, a gesture made with the hands moving, including the wrist, or a gesture that is performed with fingers moving while the hand remains fixed, were considered different. Participants largely favored gestures that made with one-hand gestures (47%) and neglected two-handed gestures (3%). With one hand, there is a slight difference between using only one finger (28%) or more fingers (22%).
- *Gesture nature:* This dimension describes the underlying meaning of a gesture, with four categories.

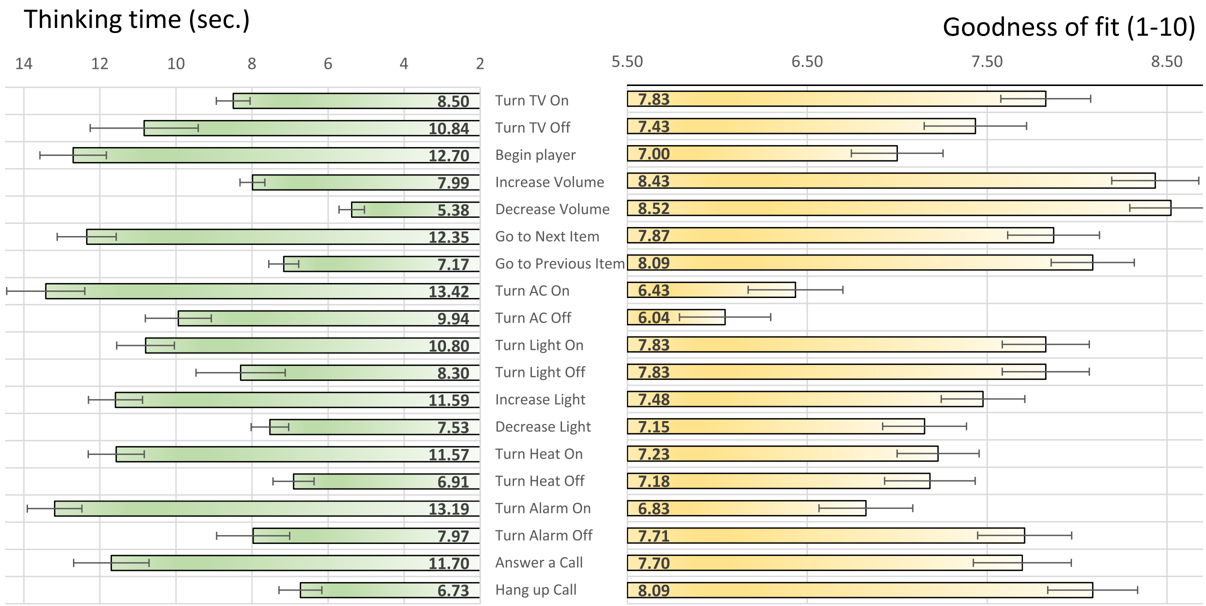


FIGURE 3. **Average thinking time (left) and goodness of fit (right) per referent. The error bars correspond to 95% confidence intervals.**

Symbolic gestures (23%) depict commonly accepted symbols employed to convey information, such as emblems and cultural gestures. This category thus represents a meaningful symbol. Metaphorical gestures (15%) are employed to shape an idea or concept, such as turning an invisible knob. Physical gestures (31%) are made when the gesture is produced as if it is physically acting on a real object. Abstract gestures (31%) are not aimed at conveying any meaning.

- *Laterality:* This dimension characterizes how the two hands are used to produce gestures, with two categories, as in many previous studies (e.g., [23], [45], [46]): (a) dominant unimanual (48%), (b) nondominant unimanual (6%), (c) symmetric bimanual (39%), and (d) asymmetric bimanual (1%). In other words, gestures performed with the left hand are considered different from the same type of gestures performed with the right hand. However, hands moving, the difference was not large. While most of the gestures made by the participants were unimanual, those that were bimanual were typically symmetric.
- *Gesture form:* This dimension specifies which form of gesture is elicited. Four categories have been distinguished: 'stroke' (when the gesture only consists of taps and flicks), 'static' (when the gesture is performed in only one location), 'static with motion' (when the gesture is performed with a static pose while the rest is moving) and 'dynamic' (when the gesture does capture any change or motion). Participants slightly favored dynamic gestures (30%) over static gestures with a micromotion (27%) and purely static gestures (22%). Strokes were almost similarly appreciated (21%), but physical strokes remain insignificant (1%).
- *Scale invariance:* This dimension determines whether gestures that are performed depend on the size, amplitude, or intensity. Overall, 64% of gestures were independent of their scale; changing their scale does not affect their meaning.

D. GENERAL ANALYSIS AND DISCUSSION

Four age groups have been formed, namely, $X<22$, $22\leq X\leq 24$, $24<X\leq 67$, and $X>67$, where X = age, based on statistics of early adopters of recent technologies. Since we had a population with various age ranges, we also tried to identify any potential correlation between age and other experimental data. The analysis focuses on potential links among the thinking time, the goodness of fit, and the creativity score that are granted to gestures and familiarity with other electronic devices, with the age variable. The implications of these findings will be further discussed in the next sections.

1) THINKING TIME

The thinking time (tt) has been averaged for all participants for each referent (Fig. 3); the thinking time ranges from $tt = 5.38$ sec for "Decrease Volume" to $tt = 13.42$ sec for "Turn AC On". Then, the elicited gestures were clustered according to dimension (Fig. 5): 1D ($M = 10.42$, $SD = 8.277$), 2D ($M = 10.43$, $SD = 8.45$), 2D1/2 ($M = 9.40$, $SD = 7.72$), and 3D ($M = 9.73$, $SD = 9.69$). To evaluate the potential normality of their respective distributions, we conducted a normality plot test; 1D (Kurtosis = 1.00 and Skewness = 1.33) and 2D1/2 (Kurtosis = 1.23 and Skewness = 1.35) satisfy the test with values within the $[-2, +2]$ range, whereas 2D and 3D do not. A Levene test was further conducted to determine whether the thinking times are normally distributed in at least one of the clusters: by means ($p = .95$),

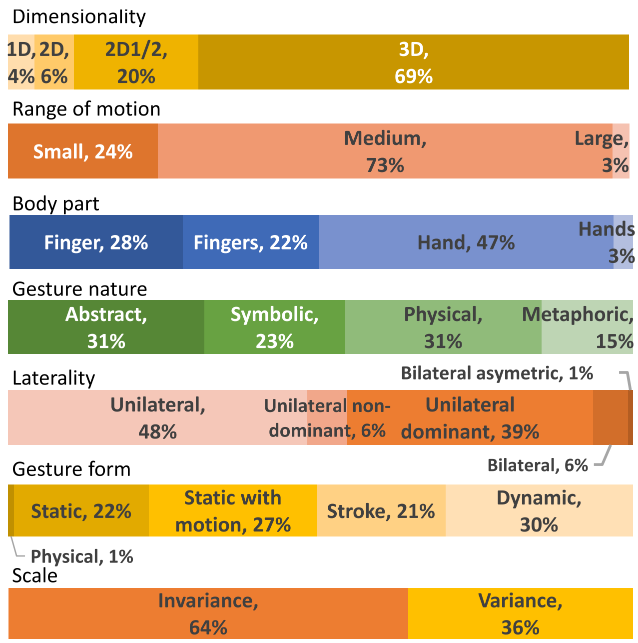


FIGURE 4. Distribution of gestures along dimensions.

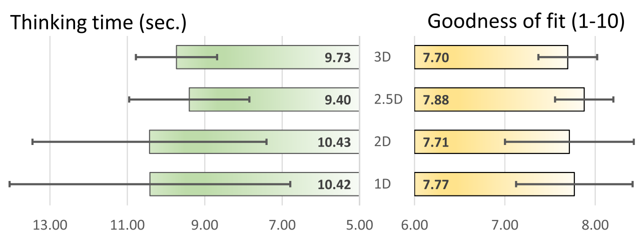


FIGURE 5. Thinking time (left) and goodness of fit (right) per gesture dimension. The error bars correspond to 95% CIs.

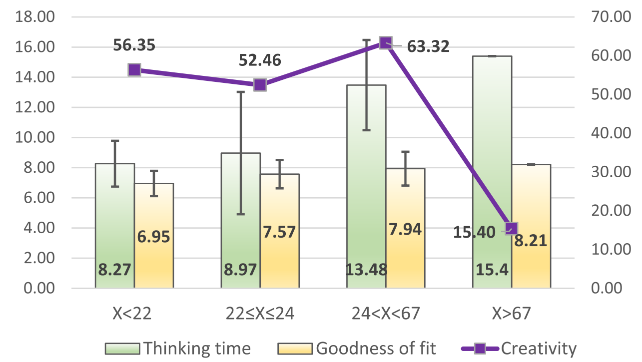


FIGURE 6. Thinking time, goodness of fit, and creativity by age group. The error bars correspond to the standard deviation.

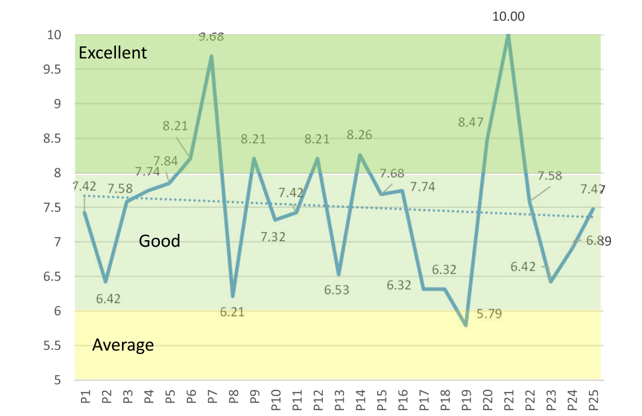


FIGURE 7. Goodness of fit per participant with 95% CI.

by medians (p = .96), and trimmed (p = .98). A single-factor ANOVA did not return any significance: SS= 34.27, df = 3, F = .13, $p > .1$, *n.s.* This result suggested that the gesture dimension did not influence the thinking time. Participants did not spend more time thinking about a spatial gesture than about a surface gesture, for instance. In terms of absolute thinking time, people took on average between 3 and 20 seconds to choose a gesture. According to Fig. 6 (green bars), younger participants chose gestures faster than older participants. The difference is quite distinct, since older participants almost took twice as much time as younger participants, namely, 8.27 seconds on average for the youngest age group and 15.40 for the oldest age group. The difference between the two first groups is subtle, however. A larger sample would be needed to confirm these results, especially for the two last categories.

2) GOODNESS OF FIT

The goodness of fit (G) has been averaged for all participants by referent (Fig. 3); it ranges from G = 6.04 for "Turn AC Off" to G = 8.52 for "Decrease Volume". Then, the elicited gestures were clustered by dimension ((Fig. 5): 1D (M = 7.77, SD = 1.19), 2D (M = 7.71, SD = 1.92), 2D1/2 (M = 7.88, SD = 1.56), and 3D (M = 7.70, SD = 1.86). To evaluate the potential normality of their respective distributions, we conducted a normality plot test; all dimensions satisfy the test, with kurtosis and skewness falling within the [−2, +2] range. However, a Levene test confirmed that the goodness of fit was normally distributed in all clusters: by means (p = .10), by medians ((p = .15), and trimmed (p = .14). A single-factor ANOVA did not return any significance: SS= 1.81, df = 3, F = .20, $p > .1$, *n.s.* This result suggests that the gesture dimension did not influence the goodness of fit. Participants were not happier when a gesture was elicited with a higher or a lower dimension. However, we discover that the goodness of the fit and the thinking time were always weakly correlated: 1D was weak (Spearman's r_s = −.11), 2D was moderate (Spearman's r_s = −.43), 2D1/2 was very weak (Spearman's r_s = −.11), and 3D was weak as well (Spearman's r_s = −.27).

During the experiment, the participants were asked to rate each gesture that they performed with a score between 0 and 10. Five intervals have been defined: very bad (for scores between 0 and 1), poor (between 1 and 3), average (between 3 and 6), good (between 6 and 8) and excellent (between 8 and 10). The average goodness of fit is graphically depicted by participant in Fig. 7 and by gesture in Fig. 8; the results indicate that people are satisfied with their

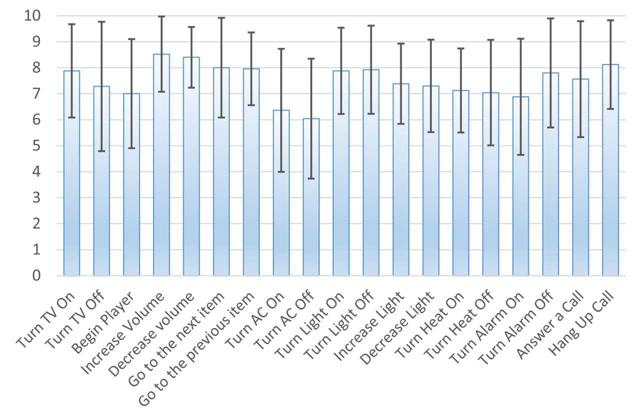


FIGURE 8. **Goodness of fit per referent with 95% confidence interval.**

gestures most of the time. Only 41 of the 475 gestures were individually assigned a goodness of fit of less than 5 (not on average). The intervals 'very bad' and 'poor' correspond to very few gestures. The gesture that is assigned the highest number of low scores (6) refers to the "Turn AC off" referent, followed by the "Turn AC on" referent (4). This result might be explained by this action being less common than answering a phone call, for example; therefore, people not easily associating a gesture with it. The "Turn TV off" and the "Start Player" referents are also assigned scores of 4. From these results, it is concluded that most gestures appeared to be natural to the participants, who regarded their choices as logical with respect to the referent.

3) SUBJECTIVE SATISFACTION

The distribution of participants' answers to the 16 questions of the IBM PSSUQ is plotted in Fig. 9, while the average score per question is plotted with a 95% confidence interval.

Beyond the neutral value of 3 for these 7-point scales, a score of 5 is the threshold beyond which the corresponding question or factor is typically considered [43]. The *system usefulness* (Q1-Q6: $M = 5.28$, $SD = 1.37$) has an overall score that satisfies this factor, with Q3 (effectiveness) raising no concern (all answers are positive) and Q5 (learning curve) raising concern. Almost all participants completed this part by answering the six questions. This was not at all the case for the second factor, namely, *information quality* (Q7-Q12: $M = 5.53$, $SD = 1.47$), for which several participants either scored the corresponding questions with disagreement or considered them not applicable (which explains why some bars are not complete in Fig. 9). The worst-scored question was Q7, which was about error messages and guidance; indeed, a device that is embedded in a smartwatch does not provide any feedback on the locus of control itself (such as the fingers and the hand) but on the device that is equipped with the radar sensor. Q8, which was about how to recover errors, elicited similar responses; it was not clear to participants how to recover an error if it occurs. Surprisingly, overall, this factor still received a score that exceeded the threshold but with the largest standard deviation, which may suggest that participants do not all agree. The third factor, namely, *interaction quality* (Q13-Q15: $M = 5.73$, $SD = 1.09$), is the most consistently assessed factor, since its score is the highest, and its deviation is the lowest among all factors. This is still reflected in the overall satisfaction (All: $M = 5.41$, $SD = 1.40$), which is still ranked above 5. These results may suggest that a contactless interaction is most appreciated with respect to its interaction capabilities but less appreciated with respect to its effectiveness and efficiency, even with the user-defined gestures.

4) EDUCATION AND CREATIVITY SCORE VARIABLES

A test was conducted to determine whether the creativity score and education are related in the experience. The education data were grouped into three categories: participants with secondary school degrees, participants with bachelor's degrees and participants with master's degrees. Participants with higher educational degrees seem to have lower creativity scores ($M = 56.68$).

However, the participants with a secondary diploma scored better on the creativity test ($M = 62.11$). We can directly link this result with the age of the participants; the younger

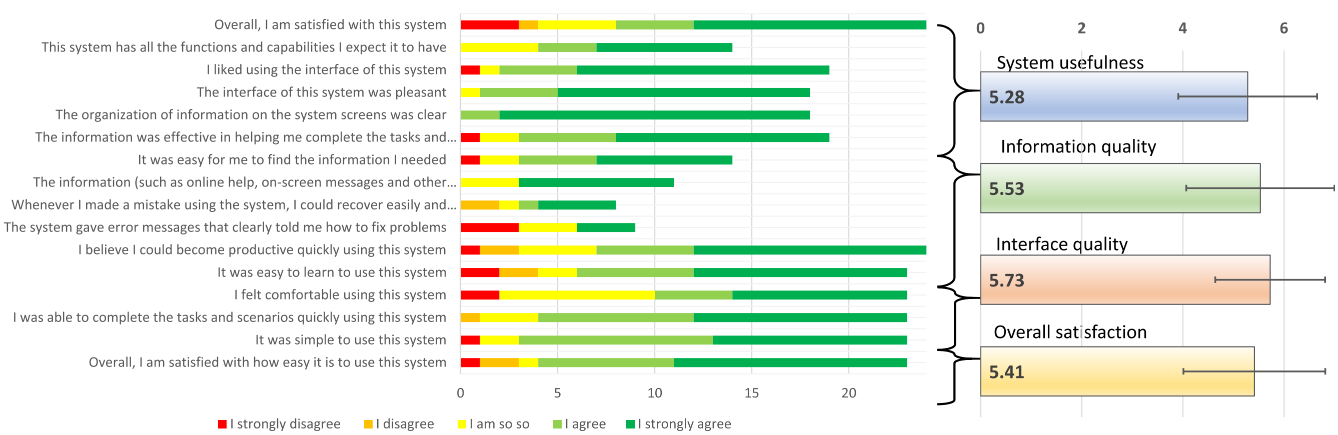


FIGURE 9. **Distribution of participants' answers to IBM PSSUQ.**

the participants are, the more creative they are regarding the use of this type of technology. We remain cautious about the results, which should be tested on a larger sample.

5) FAMILIARITY WITH OTHER DEVICES AND AGE VARIABLES

First, we focused on the correlation between the familiarity with other devices (e.g., computer, smartphone, tablet, game player, and Kinect) and the age variables. Based on 4-range means, the participants from the first group have the highest familiarity score with technological devices, and people under 22 years of age also have satisfactory scores. Based on 7-range means, the older participants have a lower familiarity rate, and people over 50 years seem to constitute the break-even point regarding the use of technologies. These results should still be tested on a larger sample of the population for generalization.

6) OTHER VARIABLES

The agreement rates of six other variables were also calculated: the body parts that are used, the laterality of the gesture, the gesture form and motion, the range of motion and the scale invariance. We selected these variables based on our appreciation of their utilities for the exploitation of radar-based sensors, such as the Google Soli [3]. The agreement rate of the range of motion that is used to perform the gesture is high, with an average value of 55.80% and a standard deviation of 11.36%. The scale invariance of the gestures also corresponded to a high agreement rate, with an average value of 52.90% and a standard deviation of 8.91%.

E. DISCUSSION OF SOME GESTURES

In this section, we will formulate hypotheses regarding the results that were presented in the previous section. The highest gesture category agreement rates were obtained for the action of turning on/off the lights.

Ten participants opened their fist (15: splay) to turn the light on and most of them chose to close their fist (14: fist) to turn the light off, which represent two symmetric gestures for two symmetric tasks. The representation of the lighting diffusion or concentration through the fingers is considered the most obvious gesture metaphor. We can also observe the use of a finger snap to turn the light on and to turn it off by 4 participants. Two of the lowest agreement rates were obtained for the action of turning on/off the phone alarm. The gesture that occurred the most often was swiping (category 10): 4 participants used this gesture to turn the alarm on and 3 to turn the alarm off. The lowest agreement rate was obtained with the "Turn AC off" action. This low agreement rate is mainly because the vast majority of our participants do not own an air conditioning system at home and, thus, have difficulty finding an appropriate gesture for an unfamiliar action.

We also calculated the agreement rates of other variables. For the range of motion, we obtained a high agreement rate of 55.8%. This high agreement rate is influenced by the small response possibility (3). However, the vast majority

FIGURE 10. Wave gesture.

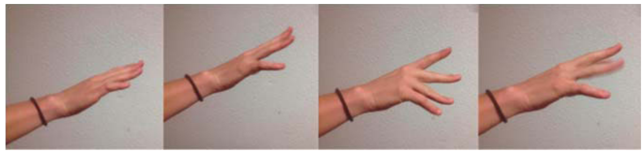


FIGURE 11. The air gesture.

of the participants chose to perform medium-sized gestures (72.8%). Almost all of the remaining participants chose to use small-sized gestures (24.0%), and only a very limited number of participants chose to use large-sized gestures (2.7%).

We found that participants typically rely on gestures that are used in daily life, for example, those that are related to the touch screens of smartphones, such as hanging up, picking up,, and browsing a list, which is another manifestation of the *legacy bias* [27]. We also had gestures that were inspired by ordinary movements, such as turning up or down the volume, turning on/off the heating and activating/deactivating the alarm. Another example of a gesture that was often used by the participants during the experiment was "Press a button of the remote control" as if they had a remote control in hand, mainly for the referent "Turn TV on/off". All of those examples seem natural, since they have become intuitive with the increasing use of current technologies in our daily lives. Nevertheless, we noticed that some participants made very special gestures, for example, the "wave" movement with the hand (Fig. 10) or with the fingers rotating in the air (Fig. 11) to turn on the air conditioning, which is a typical example of a metaphoric gesture.

F. LIMITATIONS

Although this study on radar-based sensing technologies, such as Google Soli, provide us with interesting insights, we have identified several limitations. First, the introductory video about Google Soli can influence the participants by showing how it could work for various referents. If this introduction is avoided, it can become misleading for participants who tend to issue unsupported gestures. If this introduction is included, it can become guiding but subject to a bias, even if some gestures are not directly demonstrated. Second, the pictures that are used for the referents can sway users' gestures by presenting a shape to the referent, which could vary in daily life. The representation of a referent always induces a particular behavior, since it is a single example. Changing this representation to other content (another picture) or another medium (e.g., an animation or a video) may

No.	Referent	Description
		Total of referents (30)
		Basic referents (4)
1	Turn light on	Turn on the lighting system in the room
2	Turn light off	Turn off the lighting system
3	Brighten light	Increase the luminosity level of the light
4	Dim light	Decrease the luminosity level of the light
		Selection referents (13)
5	Rate a star	Give a rating score by selecting a value from a rating bar
6	Rate a scale	Give a rating score by pointing to a value from a scale
7	Simple choice 1	Select a gender
8	Simple choice 2	Select a planet from the solar system
9	Multiple choice 1	Select many fruits
10	Multiple choice 2	Select many planets from the solar system
11	Select digital time	Select a time according to a digital format
12	Select analogic time	Select a time according to an analogic format
13	Select digital date	Select a date according to a digital format
14	Select analogic date	Select a date according to an analogic format
15	Select a season	Select one of the four seasons
16	Select a range	Select the minimum and maximum values of a range
17	Select a color	Pick a color from a palette
		Set theory referents (9)
18	Expand	Expand a set of menu items
19	Collapse	Collapse a set of menu items
20	Select	Select an item from the full set of data
21	Reselect	Select (again) from the selected subset
22	Also select	Adds a different subset to the current subset
23	Unselect	Subtracts a portion of the current subset
24	Select all	Restores the full set
25	Select none	Deactivates the full set (opposite of Select all)
26	Invert	Switches between the active and inactive portions of the set
		Specification referents (4)
27	Specify a position	Specify a physical location on a map
28	Specify a direction	Specify a routing direction on a map
29	Specify a translation	Specify a movement on the map
30	Specify a rotation	Specify a rotation of the map to align to North

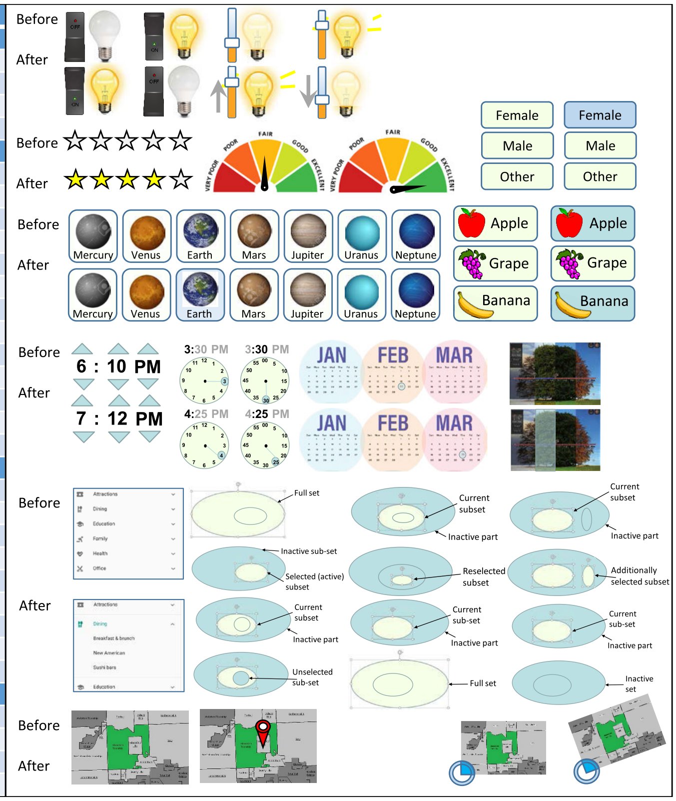


FIGURE 12. Set of thirty actions that are used in the confirmatory study and some of their corresponding referents.

influence the results. Third, the participants tend to transfer their habits with other interfaces – the most relevant example is the smartphone – without being able to think without this kind of interface. Fourth, we should reconsider the results of the creativity test. Some people scored poorly but were highly creative with the gestures, whereas others with high scores did not propose unusual gestures. Finally, in the post-study system usability questionnaire (PSSUQ), several questions were considered irrelevant in our study because gestural interaction does not provide end users with immediate feedback, guidance, or error management, unless another display is added. We remain cautious about the responses of the participants to the nonrelevant questions, which explains why several bars in Fig. 9 are incomplete.

V. CONFIRMATORY STUDY

Now that a classification of radar-based gestures is available, we would like to investigate whether participants select the correct gestures from this gesture set for a broader set of tasks. Therefore, instead of asking participants to elicit completely new gestures for referents, we asked a new sample of participants to assign a gesture from the consensus set to an action from a broader set to determine whether our consensus set is complete and correct.

A. PARTICIPANTS

Twenty voluntary participants (11 females and 9 males; aged 18 to 58 years; $M = 27.35$, $SD = 9.88$, $Mdn = 24$) were

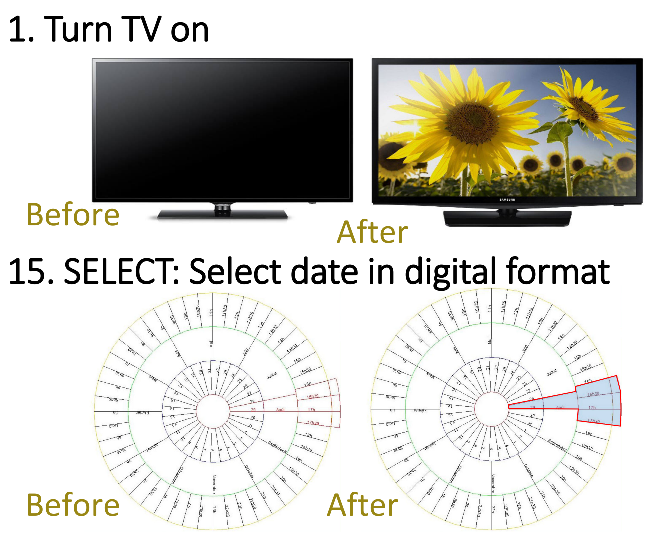


FIGURE 13. Two examples of visual priming [41]: referents from the elicitation study and from the confirmatory study.

recruited for the study via a contact list for various departments of our organization. None of these volunteers participated in the previous study. Eighteen participants were right-handed, and 2 participants were left-handed. The occupations of the participants included administrative clerk, accountant, finance specialist, teacher, psychologist, typesetter, and students in economics, law, management, and engineering. All participants reported frequent use of computers and smartphones and no dexterity problems.

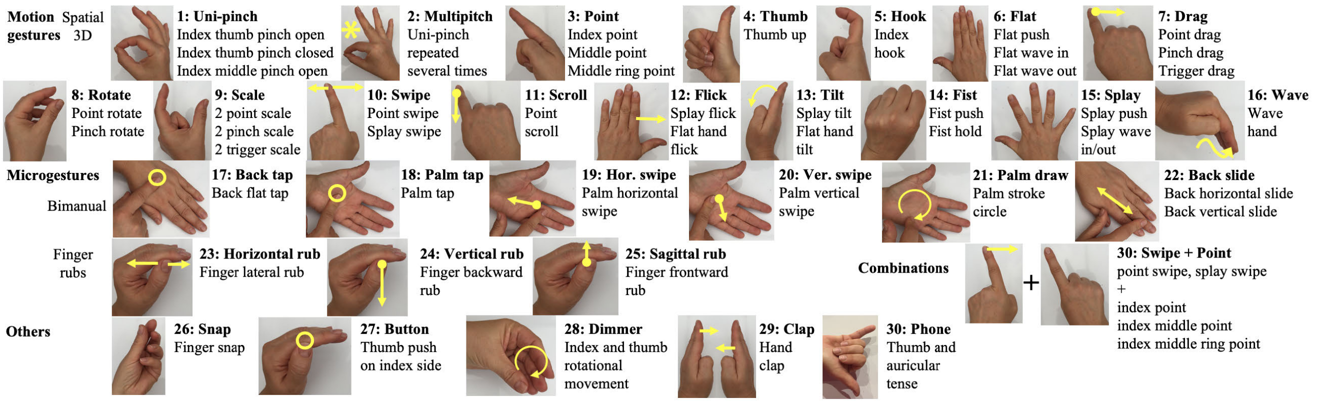


FIGURE 14. Palette of gestures from the gesture elicitation study that were presented in the confirmatory study.

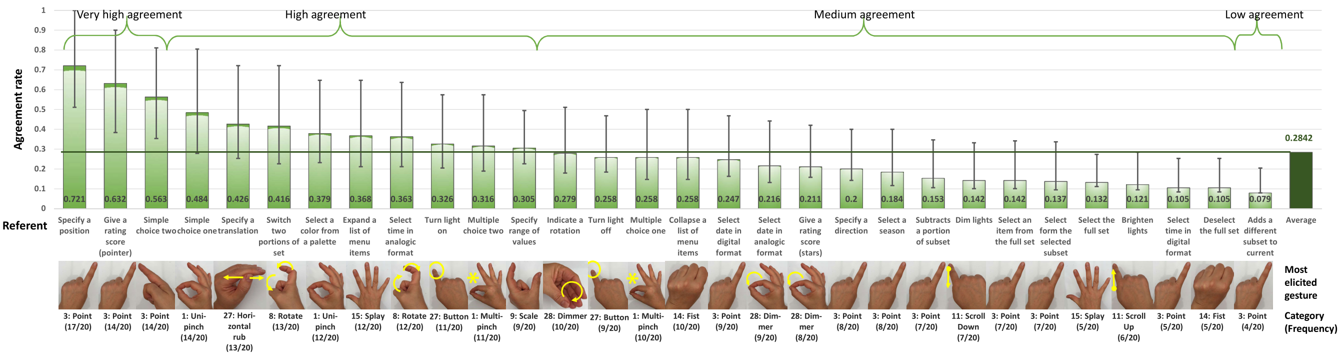


FIGURE 15. Confirmatory analysis: The agreement rates for referents, which are sorted in decreasing order with 95% confidence intervals ($\alpha = .05$).

For our study, we selected a set of 30 commands for common interactions from the current literature, which were divided into four groups, along with their referents, which were presented with a stimulus before and after the command (Fig. 12): a group of 4 basic commands (''Turn light on/off'' and ''Brighten/dim light'') [23], [47], a group of 13 *selection* commands (e.g., ''Select a date in digital/analog format'') [48], a group of 9 commands that are based on sets (e.g., ''Expand/collapse a set of menu items'', ''Select an item from the full set of data'', and ''Select (again) from the selected subset'') [48], [49], and a group of 4 referents for specification (''Specify a position, a direction, a translation, and a rotation'') [50]. These tasks have been chosen to cover an increasing number of dimensions, from 0D (e.g., ''Turn light on''), 1D (e.g., ''Dim light''), and 2D (e.g., ''Specify a position''), to 2D1/2 (e.g., ''Specify a translation'').

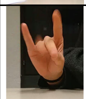
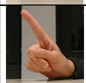
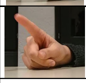
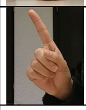
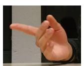
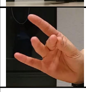
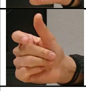
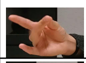
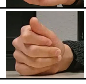
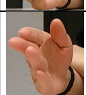
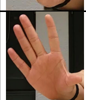
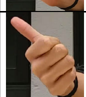
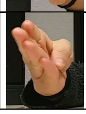
B. APPARATUS AND PROCEDURE

The experiment for this confirmatory study followed exactly the same procedure as was defined for the elicitation study (see Section III-C). The pretest and the posttest were similar, except that the procedure for the main test differed; the participants were again presented with referents that were randomly selected from the stimuli and asked to select a gesture from our classification, which was presented to them as a palette (Fig. 14). The participants were instructed to behave naturally,, to select from the palette the gesture that they thought was the most appropriate for the referent that was being presented and to specify ''another gesture'' if no appropriate gesture was identified. A gesture could be selected one or many times according to the participant's preference without any restrictions. Visual priming was used to convey the referents [41]; a simple computer screen was used to display a presentation in which the referents were randomly shown to the participants (Fig. 13 shows one referent that is being presented in the elicitation study and one in the confirmatory step with the statuses before the action and after). All gestures were recorded by a camera that was placed in front of the participants to capture their hands and fingers without capturing their faces. The measures that were computed for the gesture elicitation study were computed again for this study.

C. RESULTS AND DISCUSSION

Fig. 12 presents the agreement rates that were obtained for each referent in decreasing order, along with the corresponding most frequently elicited gesture, category, and frequency (the number of participants who selected this gesture out of 20 – if a participant selects the same gesture multiple times, it is counted as only one occurrence). All the participants

TABLE 2. Example of gesture classifications of gestures that where elicited by participant #1.

Referent	Gesture	Descriptive label and classification	Gesture category	Category number
1: Turn TV On		Thumb and index pinched together, palm oriented upwards (3D, unilateral, dynamic, physical, discrete, environment-dependent)	Thumb pinched	4
2: Turn TV Off		Thumb and middle finger pinched together, palm oriented downwards (3D, unilateral, dynamic, physical, discrete, environment-dependent)	Thumb pinched	4
3: Start player		Index pointing high, vertical swipe down (2D1/2, unilateral, static, physical, discrete, object-centric)	Point	3
4: Increase volume		Index drawing an anti-clockwise circle in the air (3D, unilateral, dynamic, physical, continuous, environment-dependent)	Rotate	8
5: Decrease volume		Index drawing a clockwise circle in the air (3D, unilateral, dynamic, symbolic, continuous, environment-dependent)	Rotate	8
6: Go to next item		Index pointing high (3D, unilateral, dynamic, physical, discrete, object-centric)	Swipe + Point	31
7: Go to previous item		Index pointing middle (3D, unilateral, dynamic, physical, discrete, object-centric)	Swipe + Point	31
8: Turn AC On		Thumb rubbing middle finger anti-clockwise (3D, unilateral, dynamic, symbolic, discrete, environment-dependent)	Palm-stroke circle	21
9: Turn AC Off		Thumb rubbing middle finger clockwise (3D, unilateral, dynamic, symbolic, discrete, environment-dependent)	Palm-stroke circle	21
10: Turn light On		Snap thumb and middle fingers, palm upwards (3D, unilateral, dynamic, symbolic, discrete, environment-dependent)	Finger snap	26
11: Turn light Off		Double snap thumb and middle fingers, palm upwards (3D, unilateral, dynamic, symbolic, discrete, environment-dependent)	Finger snap	26
12: Increase light		Thumb and middle fingers pushing against each other (3D, unilateral, dynamic, physical, continuous, environment-dependent)	Multipinch	2
13: Decrease light		Hand opening and closing as a fist (3D, unilateral, dynamic, metaphoric, continuous, environment-dependent)	Flat	6
14: Turn Heat On		Thumb rubbing against the four other fingers clockwise (3D, unilateral, dynamic, symbolic, continuous, environment-dependent)	Palm-stroke circle	21
15: Turn Heat Off		Thumb rubbing against the four other fingers anti-clockwise (3D, unilateral, dynamic, physical, continuous, environment-dependent)	Palm-stroke circle	21
16: Turn Alarm On		Opening and closing of the four major fingers (3D, unilateral, dynamic, metaphoric, discrete, environment-dependent)	Flat	6
17: Turn Alarm Off		Opening and closing hand (3D, unilateral, dynamic, metaphoric, discrete, environment-dependent)	Flat	6
18: Answer a Call		Fist with thumb moving up (2D1/2, unilateral, static, physical, discrete, environment-dependent)	Palm-stroke circle	21
19: Hang up Call		Thumb inside hand, rubbing, then exiting the palm (2D1/2, unilateral, static, physical, discrete, environment-dependent)	Palm-stroke circle	21

completed the test, but none selected "another gesture". This suggests that the participants always found an appropriate choice in the classification, which was presented as a palette, thereby fostering the completeness of the classification.

Overall, the agreement rates, which were again computed by AGATe [20], are much higher in magnitude in this confirmatory study than in the elicitation study: the referents were ranked between .079 and .721 (M = .284,

$SD = .016$). We conducted a one-tailed Mann-Whitney test for the two independent samples. There is a very highly significant difference in the rates between the elicitation and confirmatory studies ($U = 67$, $M = 315$, $SD = 52.23$, $Z = 4.74$, $p^{***} < .001$) with a significant effect of $r = .66$, which suggests that higher agreement or convergence across participants is realized when they select gestures from the classification than when they elicit gestures for themselves.

Three referents realized very high agreement rates: "Specify a position" ($AR = .721$), "Give a rating score" ($AR = .632$), and "Simple choice two" ($AR = .563$). Nine referents out of the set of thirty achieved high agreement rates ($.5 < AR < .3$), the vast majority achieved moderate agreement rates ($N = \frac{17}{30} = 57\%, .3 \leq AR < .1$), and only one referent, which is a difficult operation, had a low agreement rate ("Add a different subset"). Participants correctly selected the "3:Button" category for 0D tasks (e.g., "Turn light On/Off"). In 1D tasks, participants tend to prefer the "3:Point" category, followed by uni- and multipinch gestures, which correspond to the "28:Dimmer" category. "7: Horizontal rub" was also suitably selected for specifying a translation and "9:Scale" for specifying a range of values.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a gesture elicitation study that explores the contactless gestures that are preferred by end users with a radar sensing technology (Google Soli). We created our own design space from participants' gestures and analyzed the results of each subjective evaluation. We identified gestures that refer to daily activities in a home. By describing a procedure for classifying gestures, there was consensus among participants for common gestures, such as "Dimmer", "Virtual remote control" and "Phone", while many gestures are based on users' classical preferences that we classified. The other gestures utilized more innovative interactions techniques, which also caught our attention. We believe that hand gesture recognition is a promising approach for facilitating the lives of people with reduced mobility and people with a busy lifestyle. However, we observed that the participants were noticeably biased by pictures of the referents. Based on these results, it seems that the participants chose the point or swipe gestures more frequently than other gestures.

Furthermore, most participants have some digital knowledge due to their smartphone or their laptop. In the future, it could be interesting to determine whether these users' preferences are observed among a larger sample of people. Interestingly, we observed some older participants transmitting gestures from old-fashioned lamps or phones. It seems that this generation is transferring their gestures from older devices. A more advanced study could analyze individuals in their own daily environments and evaluate the correlation with our results. We carried out our study in a quiet and controlled environment; however, participants are supposed to use this prototype in their daily lives in various contexts of use. For example, a 'standard' living room could be a more suitable setting. Therefore, we plan to conduct another elicitation study *in vivo* as opposed to *in vitro* and to determine the extent to which another set of referents may influence the results. It is likely that the resulting classification (Table 1) will remain constant over time, as suggested by the confirmatory study. Although a few gestures may always appear, it is likely that the agreement scores and their corresponding preferred gestures may vary depending on users and their tasks.

Future work should focus on testing various gesture recognition algorithms on this gesture set to determine the confusion matrix and the extent to which the distance between the sensor and the gesture may affect the signal intensity and the recognition accuracy, as in [3]. A representative example of this kind of study is [8], which demonstrates that the lift microgesture group has lower error rates than the pinch group. Microgestures were sorted in increasing order of the error rate: index lift ($\epsilon < .05\%$), middle lift ($\epsilon < .1\%$), ring lift and index pinch ($\epsilon < .15\%$), and middle and pinky pinches ($\epsilon < .2\%$).

APPENDIX A
GESTURE CLASSIFICATION

See Table 2 here.

REFERENCES

[1] C. J. Hansen, "WiGig: Multi-gigabit wireless communications in the 60 GHz band," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 6–7, Dec. 2011.

[2] H.-S. Yeo, G. Flamich, P. Schrempf, D. Harris-Birtill, and A. Quigley, "RadarCat: Radar categorization for input & interaction," in *Proc. ACM 29th Annu. Symp. User Interface Softw. Technol. (UIST)*, New York, NY, USA, 2016, pp. 833–841, doi: 10.1145/2984511.2984515.

[3] S. Wang, J. Song, J. Lien, I. Poupyrev, and O. Hilliges, "Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum," in *Proc. ACM 29th Annu. Symp. User Interface Softw. Technol. (UIST)*, New York, NY, USA, 2016, pp. 851–860, doi: 10.1145/2984511.2984565.

[4] L. Gang, S. Zhang, F. Fioranelli, and H. Griffiths, "Effect of sparsity-aware time-frequency analysis on dynamic hand gesture classification with radar micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 12, no. 8, pp. 815–820, Apr. 2018.

[5] G. Li, R. Zhang, M. Ritchie, and H. Griffiths, "Sparsity-driven micro-Doppler feature extraction for dynamic hand gesture recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 2, pp. 655–665, Apr. 2018.

[6] Z. Zhang, Z. Tian, Z. Mu, and Y. Liu, "Application of FMCW radar for dynamic continuous hand gesture recognition," in *Proc. Int. Workshop Millim.-Wave Radar Technol. Appl.*, Sep. 2018, pp. 298–303.

[7] J. O. Wobbrock, M. R. Morris, and A. D. Wilson, "User-defined gestures for surface computing," in *Proc. ACM Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2009, pp. 1083–1092, doi: 10.1145/1518701.1518866.

[8] D. Way and J. Paradiso, "A usability user study concerning free-hand microgesture and wrist-worn sensors," in *Proc. 11th Int. Conf. Wearable Implant. Body Sensor Netw.*, Jun. 2014, pp. 138–142.

[9] E. Chan, T. Seyed, W. Stuerzlinger, X.-D. Yang, and F. Maurer, "User elicitation on single-hand microgestures," in *Proc. ACM Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2016, pp. 3403–3414, doi: 10.1145/2858036.2858589.

[10] W. Wang, A. X. Liu, and K. Sun, "Device-free gesture tracking using acoustic signals: Demo," in *Proc. ACM 22nd Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, New York, NY, USA, 2016, pp. 497–498, doi: 10.1145/2973750.2987385.

[11] S. Yun, Y.-C. Chen, H. Zheng, L. Qiu, and W. Mao, "Strata: Fine-grained acoustic-based device-free tracking," in *Proc. ACM 15th Annu. Int. Conf. Mobile Syst., Appl., Services (MobiSys)*, New York, NY, USA, 2017, pp. 15–28, doi: 10.1145/3081333.3081356.

[12] T. Babic, H. Reiterer, and M. Haller, "GestureDrawer: One-handed interaction technique for spatial user-defined imaginary interfaces," in *Proc. ACM 5th Symp. Spatial User Interact. (SUI)*, New York, NY, USA, 2017, pp. 128–137, doi: 10.1145/3131277.3132185.

[13] P. I. C. Martinez, S. De Pirro, C. T. Vi, and S. Subramanian, "Agency in mid-air interfaces," in *Proc. ACM CHI Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2017, pp. 2426–2439, doi: 10.1145/3025453.3025457.

[14] M. I. Skolnik, *Introduction to Radar Systems*. New York, NY, USA: McGraw-Hill, 2001.

[15] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 142-1–142-19, Jul. 2016, doi: 10.1145/2897824.2925953.

[16] H.-S. Yeo and A. Quigley, "Radar sensing in human-computer interaction," *Interactions*, vol. 25, no. 1, pp. 70–73, Dec. 2017, doi: 10.1145/3159651.

[17] F. Bernardo, N. Arner, and P. Batchelor, "O Soli Mio: Exploring millimeter wave radar for musical interaction," in *Proc. 17th Int. Conf. New Inter. Musical Expression (NIME)*. Aalborg, Denmark: Aalborg Univ., May 2017, pp. 283–286.

[18] B. Ens, A. Quigley, H.-S. Yeo, P. Irani, and M. Billinghurst, "Multi-scale gestural interaction for augmented reality," in *Proc. ACM SIGGRAPH Asia Mobile Graph. Interact. Appl. (SA)*, New York, NY, USA, 2017, pp. 11-1–11-2, doi: 10.1145/3132787.3132808.

[19] Z. Zhang, Z. Tian, and M. Zhou, "Latern: Dynamic continuous hand gesture recognition using FMCW radar sensor," *IEEE Sensors J.*, vol. 18, no. 8, pp. 3278–3289, Feb. 2018.

[20] R.-D. Vatavu and J. O. Wobbrock, "Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit," in *Proc. 33rd ACM Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2015, pp. 1325–1334, doi: 10.1145/2702123.2702223.

[21] R.-D. Vatavu and J. O. Wobbrock, "Between-subjects elicitation studies: Formalization and tool support," in *Proc. ACM CHI*. New York, NY, USA, 2016, pp. 3390–3402, doi: 10.1145/2858036.2858228.

[22] J. Ruiz, Y. Li, and E. Lank, "User-defined motion gestures for mobile interaction," in *Proc. ACM SIGCHI Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2011, pp. 197–206, doi: 10.1145/1978942.1978971.

[23] R.-D. Vatavu, "User-defined gestures for free-hand TV control," in *Proc. ACM 10th Eur. Conf. Interact. TV Video (EuroITV)*, New York, NY, USA, 2012, pp. 45–48, doi: 10.1145/2325616.2325626.

[24] R.-D. Vatavu and I.-A. Zaiţi, "Leap gestures for TV: Insights from an elicitation study," in *Proc. ACM Int. Conf. Interact. Exper. TV Online Video (TVX)*, 2014, pp. 131–138, doi: 10.1145/2602299.2602316.

[25] G. Calvary, J. Coutaz, D. Thevenin, Q. Limbourg, L. Bouillon, and J. Vanderdonckt, "A unifying reference framework for multi-target user interfaces," *Interacting Comput.*, vol. 15, no. 3, pp. 289–308, 2003, doi: 10.1016/S0953-5438(03)00010-9.

[26] P. Koutsabasis and P. Vogiatzidakis, "Empirical research in mid-air interaction: A systematic review," *Int. J. Hum. Comput. Interact.*, vol. 35, no. 18, pp. 1747–1768, 2019, doi: 10.1080/10447318.2019.1572352.

[27] M. R. Morris, J. O. Wobbrock, and A. D. Wilson, "Understanding users' preferences for surface gestures," in *Proc. Graph. Interface (GI)*, Toronto, ON, Canada: Canadian Information Processing Society, 2010, pp. 261–268. [Online]. Available: http://dl.acm.org/citation.cfm?id=1839214.1839260

[28] H. Dong, A. Danesh, N. Figueroa, and A. El Saddik, "An elicitation study on gesture preferences and memorability toward a practical hand-gesture vocabulary for smart televisions," *IEEE Access*, vol. 3, pp. 543–555, 2015.

[29] M. Berthellemy, E. Cayez, M. Ajem, G. Bailly, S. Malacria, and E. Lecolinet, "SpotPad, LociPad, ChordPad and InOutPad: Investigating gesture-based input on touchpad," in *Proc. ACM 27th Conf. L'Interact. Homme-Mach. (IHM)*, New York, NY, USA, 2015, pp. 4:1–4:8, doi: 10.1145/2820619.2820623.

[30] D. Chattopadhyay and D. Bolchini, "Motor-intuitive interactions based on image schemas: Aligning touchless interaction primitives with human sensorimotor abilities," *Interacting Comput.*, vol. 27, no. 3, pp. 327–343, 2015.

[31] B.-F. Gheran, J. Vanderdonckt, and R.-D. Vatavu, "Gestures for smart rings: Empirical results, insights, and design implications," in *Proc. ACM Designing Interact. Syst. Conf. (DIS)*, New York, NY, USA, 2018, pp. 623–635, doi: 10.1145/3196709.3196741.

[32] T. Döring, D. Kern, P. Marshall, M. Pfeiffer, J. Schöning, V. Gruhn, and A. Schmidt, "Gestural interaction on the steering wheel: Reducing the visual demand," in *Proc. ACM SIGCHI Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2011, pp. 483–492, doi: 10.1145/1978942.1979010.

[33] I. B. Rodriguez and N. Marquardt, "Gesture elicitation study on how to opt-in & opt-out from interactions with public displays," in *Proc. ACM Int. Conf. Interact. Surf. Spaces (ISS)*, New York, NY, USA, 2017, pp. 32–41, doi: 10.1145/3132272.3134118.

[34] I. Gatto and F. Pittarello, "Prototyping a gestural interface for selecting and buying goods in a public environment," in *Proc. ACM Int. Work. Conf. Adv. Vis. Interfaces (AVI)*, New York, NY, USA, 2012, pp. 784–785, doi: 10.1145/2254556.2254713.

[35] I. Akpan, P. Marshall, J. Bird, and D. Harrison, "Exploring the effects of space and place on engagement with an interactive installation," in *Proc. ACM SIGCHI Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2013, pp. 2213–2222, doi: 10.1145/2470654.2481306.

[36] D. Mauney, J. Howarth, A. Wirtanen, and M. Capra, "Cultural similarities and differences in user-defined gestures for touchscreen user interfaces," in *Proc. CHI ACM Extended Abstracts Hum. Factors Comput. Syst. (CHI EA)*, New York, NY, USA, 2010, pp. 4015–4020, doi: 10.1145/1753846.1754095.

[37] I. Bostan, O. T. Buruk, M. Canat, M. O. Tezcan, C. Yurdakul, T. Göksun, and O. Özcan, "Hands as a controller: User preferences for hand specific on-skin gestures," in *Proc. ACM Conf. Designing Interact. Syst. (DIS)*, New York, NY, USA, 2017, pp. 1123–1134, doi: 10.1145/3064663.3064766.

[38] H. Havlucu, M. Y. Ergin, I. Bostan, O. T. Buruk, T. Göksun, and O. Özcan, "It made more sense: Comparison of user-elicited on-skin touch and freehand gesture sets," in *Distributed, Ambient and Pervasive Interactions*, N. Streitz and P. Markopoulos, Eds. Berlin, Germany: Springer, 2017, pp. 159–171.

[39] M. Korkman, U. Kirk, and S. Kemp, *A Developmental Neuropsychological Assessment Manual*. San Antonio, TX, USA: Psychological Corporation, 1998.

[40] M. R. Morris, A. Danielescu, S. Drucker, D. Fisher, B. Lee, M. C. Schraefel, and J. O. Wobbrock, "Reducing legacy bias in gesture elicitation studies," *Interactions*, vol. 21, no. 3, pp. 40–45, May 2014, doi: 10.1145/2591689.

[41] F. Cafaro, L. Lyons, and A. N. Antle, "Framed guessability: Improving the discoverability of gestures and body movements for full-body interaction," in *Proc. ACM CHI Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2018, pp. 593-1–593-12, doi: 10.1145/3173574.3174167.

[42] J. R. Lewis, "Psychometric evaluation of the PSSUQ using data from five years of usability studies," *Int. J. Hum.-Comput. Interact.*, vol. 14, nos. 3–4, pp. 463–488, 2002, doi: 10.1080/10447318.2002.9669130.

[43] J. R. Lewis, "Sample sizes for usability tests: Mostly math, not magic," *Interactions*, vol. 13, no. 6, pp. 29–33, Nov. 2006, doi: 10.1145/1167948.1167973.

[44] T. Seyed, C. Burns, M. C. Sousa, F. Maurer, and A. Tang, "Eliciting usable gestures for multi-display environments," in *Proc. ACM Int. Conf. Interact. Tabletops Surf. (ITS)*, New York, NY, USA, 2012, pp. 41–50, doi: 10.1145/2396636.2396643.

[45] C. Kühnel, T. Westermann, F. Hemmert, S. Kratz, A. Müller, and S. Möller, "I'm home: Defining and evaluating a gesture set for smart-home control," *Int. J. Hum.-Comput. Stud.*, vol. 69, no. 11, pp. 693–704, 2011, doi: 10.1016/j.ijhcs.2011.04.005.

[46] H.-N. Liang, C. Williams, M. Semegen, W. Stuerzlinger, and P. Irani, "User-defined surface+motion gestures for 3D manipulation of objects at a distance through a mobile device," in *Proc. ACM 10th Asia–Pacific Conf. Comput. Hum. Interact. (APCHI)*, New York, NY, USA, 2012, pp. 299–308, doi: 10.1145/2350046.2350098.

[47] J. Bobeth, S. Schmehl, E. Kruijff, S. Deutsch, and M. Tscheligi, "Evaluating performance and acceptance of older adults using freehand gestures for TV menu control," in *Proc. ACM EuroITV*. New York, NY, USA, 2012, pp. 35–44, doi: 10.1145/2325616.2325625.

[48] J. M. González-Calleros, J. Guerrero-García, J. Vanderdonckt, and J. Munoz-Arteaga, "Towards canonical task types for user interface design," in *Proc. Latin Amer. Web Congr., Joint LA-WEB/CLIHC Conf.*, Yucatan, Mexico, Nov. 2009, pp. 63–70, doi: 10.1109/LA-WEB.2009.33.
[49] I.-A. Zaiţi, Ş. G. Pentiuc, and R.-D. Vatavu, "On free-hand TV control: Experimental results on user-elicited gestures with leap motion," *Pers. Ubiquit. Comput.*, vol. 19, nos. 5–6, pp. 821–838, 2015.
[50] J. D. Foley, V. L. Wallace, and P. Chan, "The human factors of computer graphics interaction techniques," *IEEE Comput. Graph. Appl.*, vol. 4, no. 11, pp. 13–48, Nov. 1984.

NATHAN MAGROFUOCO received the master's degree *(cum laude)* in computer science, and the master's degree in management sciences *(magna cum laude)* from the Université Catholique de Louvain (UCLouvain), Belgium, in 2016 and 2017, respectively, He is currently pursuing the Ph.D. degree in computer science with the Louvain Research Institute in Management and Organizations (LouRIM), UCLouvain. His research interests include gesture interaction and multicontext stroke gesture recognition based on Clifford's algebra.

JORGE-LUIS PÉREZ-MEDINA received the M.S. degree in information systems and the Ph.D. degree in computer science from the Joseph Fourier University of Grenoble, France, in 2006 and 2010, respectively. He was a Postdoctoral Researcher in human-computer interaction with the Université Catholique de Louvain, Belgium, from 2015 to 2016. He is currently a Professor and a Researcher with the Intelligent and Interactive Systems Laboratory (SI2 Lab), Universidad de Lasméricas, Quito, Ecuador. His research interests include facilitating the work of IT managers and model designers by helping them select methods, and models and modeling environments that are adapted to their needs.

PAOLO ROSELLI has been a Senior Research Associate with tenure (Ricercatore universitario, confermato) in mathematical analysis with Università di Roma "Tor Vergata", Italy, since 2001. He was an Invited Professor with the Université Catholique de Louvain, from 2012 to 2020 and the Université Libre de Bruxelles, Belgium from 2014 to 2015 and 2018 to 2019. He has long experience in graduate and undergraduate courses of Linear Algebra, Calculus, Basic and advanced mathematical analysis, measure theory, geometric algebras for students in mathematics, physics, engineering, computer science, and natural sciences. His research interests include variational methods in Hamiltonian systems, critical point theory, symmetry and symmetry-breaking phenomena, real analysis and fine properties in Lebesgue function spaces, logic, Curry-Howard correspondence and didactics, applications of Clifford geometric algebras to geometric measures, integration, and gesture recognition. He received the prize of the Royal Academy of Sciences, Letters and Fine Arts of Belgium, in 2014.

JEAN VANDERDONCKT (M'04–SM'08) received the M.Sc. degree in mathematics, the M.Sc. degree in computer science, and the Ph.D. degree from the University of Namur, Belgium, in 1987, 1989, and 1997, respectively. He was a Visiting Associate Professor with Stanford University, in 2000, and permanent Invited Professor with the Polytechnic University of Valencia. He is currently a Full Professor with the Louvain School of Management, Université Catholique de Louvain, Belgium, where he has been led the Louvain Interaction Laboratory, since 1998. His research interests include human–computer interaction (HCI), engineering interactive computing systems (EICS), intelligent user interfaces (IUI), usability engineering and software engineering. He is also an Associate Editor of *ACM Transactions on Interactive Intelligent Systems* (TiiS) and *Knowledge and Information Systems* (KAIS). He is also the Co-Editor-in-Chief of the Series of *Human–Computer Interaction* (Springer) and the Briefs in *Human–Computer Interaction* (Springer). He is the Conference Co-Chair of IEEE RCIS'2019 and a Full Paper Co-Chair of ACM EICS'2020.

SANTIAGO VILLARREAL received the bachelor's degree in computer science and the master's degree in web intelligence from Jean Monnet University, Saint-Étienne, France, in 2007 and 2010, respectively. He is currently pursuing the Ph.D. degree in computer science with UCLouvain. He was a Professor with the Universidad de Las Américas, where he is teaching software development. He has also collaborated with the Intelligent and Interactive Systems Laboratory (SI2 Lab-UDLA) on the ePHoRt project. He is also a full-time Researcher who is funded by research convention n°7901 (Emerald Cast Assistant, MecaTech competitivity pole, DGO6). His research interests include human–computer interaction, gesture elicitation studies, artificial intelligence applications in movement, multiagent systems, the distributed constraint satisfaction problem (DisCSP), and human movement processing.