Engineering Touchscreen Input for 3-Way Displays: Taxonomy, Datasets, and Classification

Jean Vanderdonckt
Université catholique de Louvain
Louvain Research Institute in Management and Organizations
Louvain-la-Neuve, Belgium
jean.vanderdonckt@uclouvain.be

Radu-Daniel Vatavu
MintViz Lab, MANSiD Center
Ştefan cel Mare University of Suceava
Suceava, Romania
radu.vatavu@usm.ro

Arthur Sluÿters
Université catholique de Louvain
Louvain Research Institute in Management and Organizations
Louvain-la-Neuve, Belgium
arthur.sluyters@uclouvain.be

Figure 1: A 3-way display is a reconfigurable form factor of a personal computer that features one central display and two symmetrically expandable lateral sides. In this work, we examine user-defined gesture input for personal 3-way display devices.

ABSTRACT

In the family of personal multi-display devices and environments, 3-way displays conveniently integrate into the conventional form factors of laptops and tablets, featuring both a central display area and two symmetrically expandable lateral sides. However, despite a large body of knowledge on touch input for single-display devices, little is known about users' gesture preferences for 3-way displays. We propose a cross-display gesture taxonomy for future explorations of gesture input for multi-display devices, in which we position 3-way displays. Using a requirement elicitation, we report results from two gesture elicitation studies with a total of 48 participants, where a 3-way display was used as a remote control panel for a smart home environment (study #1) and a touchscreen interface for content manipulation performed both within and across displays (study #2). Based on these findings, we offer two consensus datasets of 3-way-display gestures that are consolidated into a larger classification of stroke-gesture input for 3-way displays.

EICS Companion '24, June 24–28, 2024, Cagliari, Italy

ACM ISBN 979-8-4007-0651-6/24/06
https://doi.org/10.1145/3660515.3661331

CCS CONCEPTS

• **Software and its engineering → Requirements analysis**; • **Hardware → Touch screens**; • **Human-centered computing → Gestural input**; *User studies*; Laboratory experiments; *Participatory design*.

KEYWORDS

Gesture-based User Interfaces, Multi-display devices, New datasets, Requirements elicitation, Requirements engineering, Stroke gestures, Touch input, Three-way displays.

ACM Reference Format:
Jean Vanderdonckt, Radu-Daniel Vatavu, and Arthur Sluÿters. 2024. Engineering Touchscreen Input for 3-Way Displays: Taxonomy, Datasets, and Classification. In *Companion of the 16th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS Companion '24), June 24–28, 2024, Cagliari, Italy*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3660515.3661331

1 INTRODUCTION

Multi-Display Environments (MDEs) distribute interactions over multiple displays with form factors and contexts of use spanning from personal mobile devices intended for individual users [52] to multitouch tabletop computing [8] and wall displays suited for collaborative tasks involving multiple users [22,32]. MDEs also include touchscreen displays that customize interactions according to display size [2], position [31], resolution [2], and mobility [38] with many benefits regarding user experience [10]. However, due to the distribution of the display usage in both space and time, MDEs pose challenges in synchronizing individual displays [15], initiating communication [17], transferring applications [7] and data [16] across displays, and manipulating content [15,17,22], respectively, in smart and ambient intelligence environments, including smart rooms and smart buildings. To address such aspects, gesture input has been examined for interactions with MDEs [32,33], including on-screen [55] and motion-based [52] gestures.

Personal MDEs, in the form of multi-display devices (MDDs), take advantage of reconfigurable form factors in both mobile and stationary contexts of use. For example, 2-way display smartphones, such as Samsung Galaxy Fold, enable various input modalities, including flipping gestures [52], while 3-way monitor configurations, e.g., Teamgee Triple Monitor Screen Extender, offer customized content visualization and effective multitasking. However, while input for large-scale or multi-device MDEs has been extensively examined before [7,8,13–17,22,32,33], little is known about users' gesture input preferences for personal MDDs, such as the ones in the above examples, and about how MDDs with distinctive form factors can be used to effect commands in a smart environment.

To engineer gesture-based user interfaces for interactive applications using a 3-way display, we need an appropriate gesture vocabulary [51] and an efficient recognizer [1]. While the second problem of recognizing stroke gestures on a touch surface has been largely resolved thanks to the availability [5] of compact and efficient recognizers [9], such as template-based ones [19], the first problem of the gesture vocabulary remains open and unsolved.

To address this problem, we contribute conceptual and empirical findings for touchscreen input on 3-way displays. We are specifically interested in the 3-way form factor since it conveniently packs multiple displays into conventional computer monitors, laptops, or tablets, featuring both a central display area and two symmetrically expandable lateral sides (see Fig. 1). Regarding requirements engineering in general and elicitation in particular, a popular method for uncovering end users' preferences for gesture input consists of conducting a Gesture Elicitation Study (GES) [50], a participatory design method where participants are instructed to propose one or many gestures in response to a referent materializing an action or a task (see [45] for a systematic literature review). Using this method, our contributions are manifold:

- Section 2 revisits Brudy *et al.*'s [4] cross-device taxonomy of interactions in the light of gesture input spanning multiple displays to propose a new cross-display gesture taxonomy, in which 3-way displays are positioned.
- Section 3 reports a first GES using the 3-way display as a generic remote control panel for a smart environment, uncovering end users' preferences for taps and directional swipes performed using the center and right displays, predilection for input centrality and following participants' handedness.
- Section 4 reports a second GES for content manipulation performed both within and across the displays of the 3-way display for effecting commands in a smart environment.
- Section 5 summarises the gesture datasets resulting from the two GES, consolidates them into a classification of 3-way stroke gestures, and discusses these results.

2 RELATED WORK

Nacenta et al. [22] distinguish between spatially-aware and spatially-agnostic input techniques in their taxonomy of cross-display object movement in MDEs. In the former case, displays are referenced in an absolute manner, according to their spatial location in the MDE [6,27]. In the latter, they are referenced in a relative manner through specific techniques, such as physical identifiers, lists, hierarchies, and coding schemes [27]. Due to such differences in how individual displays can be referenced in MDEs, we expect differences to also exist in users' mental models for gesture-based input to interact with personal MDDs. To structure related work in this area, we revisit Brudy *et al.*'s [4] taxonomy for cross-device interaction with a focus on gesture input for MDEs/MDDs as follows:

(a) *Multi-monitor/screen systems* foster 2-way display gestures, such as flipping [52], pulling [26], and gestures performed across and between screens [26], respectively. For example, Yang et al. [52] proposed a design space of thirty gestures consisting of flipping actions for 2-way display smartphones. They found that flipping gestures performed with the wrist were the fastest, while bimanual gestures were the most preferred by the participants in their study. Shen and Harrison [34] proposed a design space of pull gestures for 2-way display laptops, structured according to the location of interaction (on screen *vs.* off-screen) and the number of screens (one *vs.* two). Examples of on-screen gestures include drag, flick, pinch-to-zoom, double tap, click & hold, lasso, pull apart, dial, and knuckle taps, while cross-screen input is mainly represented by drag and drag & tap, respectively.
(b) *Multi-slate/tablet systems* afford both spatially-aware and spatially-agnostic gesture input [6,27], shortcuts [25], motion input [30], stitching [14], and dexterous finger-based gestures [53]. For example, Hinckley *et al.*'s [14] stitching technique enables pen-based gesture input to be performed across multiple displays.
(c) *Cross-display interactions* involve object movement gestures [22] performed across multiple displays. For example, a GES conducted for mobile cross-display tasks by Rädle et al. [27] revealed a percentage of 71% of the elicited gestures to be spatially aware. Overall, spatially-aware gestures are generally preferred to spatially-agnostic ones as they present lower mental demand, effort, and frustration [6].
(d) *Cross-surface gesture interactions* are mostly represented by the pick & drop technique [29], which implements copy/-paste and content transfer across surfaces [49] and objects [40]. Other examples include multitouch [8], swipe-hand open [26], prevalent gestures [32], mono-surface [21,50], and multi-surface gestures [33].
(e) *Cross-device interactions* include gestures for data sharing [7] and transfer [16], connection gestures [15,17], auxiliary display input [26], and cross-device drag & drop [35]. Also, synchronous gestures [13] enable individual users to perform the same gesture on two devices to link them or two users to perform the gesture simultaneously on their own devices and achieve the same result. In the context of MDEs, synchronous gestures have been applied to cross-device input [28] to implement connection-action phrases involving one-to-one

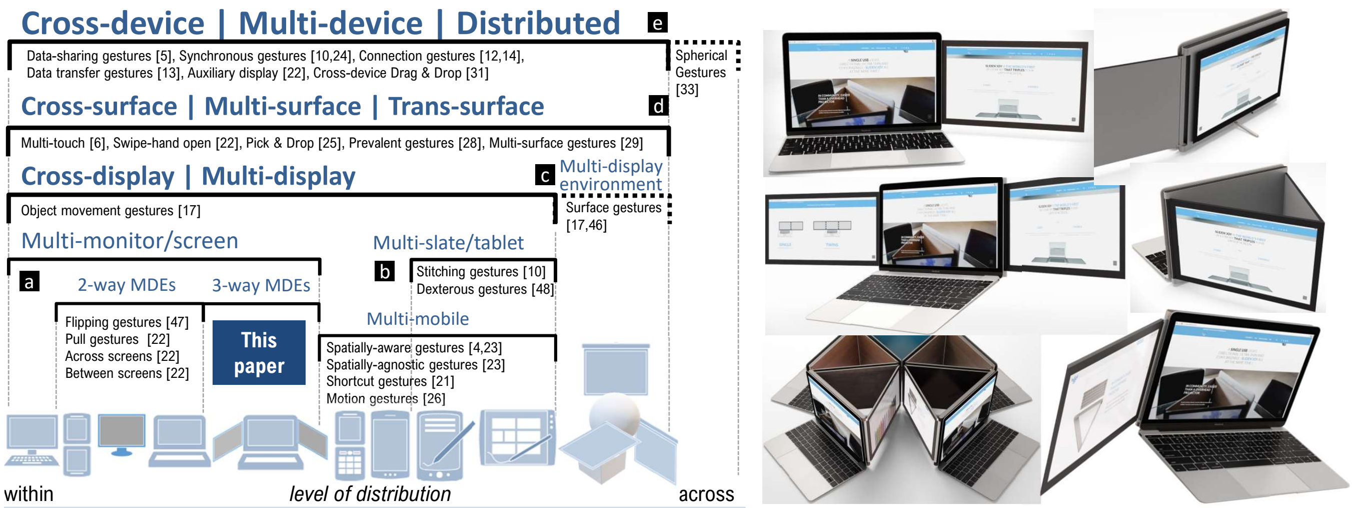


Figure 2: *Left:* MDE/MDD gesture taxonomy, based on Brudy et al. [4]'s taxonomy. *Right:* various non-flat arrangements enabled by the 3-way display used in our studies, to be leveraged in future work on combined touch and manipulative gesture input.

and one-to-many associations with flexible physical arrangements of devices. Kray et al. [15] examined phone-to-phone, phone-to-tabletop, and phone-to-public display cross-device input in a gesture elicitation study conducted to collect users' preferences for smartphone gestures. In the same area, Seyed et al. [32] conducted a gesture elicitation study to identify suitable commands for content transfer across tablets, tabletops, and wall displays, and reported that directional swipes were preferred by users. Finally, Soni et al. [37] used gesture elicitation to understand the characteristics of gesture input articulated on a spherical device. Their results revealed users are likely to perform multi-finger and whole-handed input on the spherical display than on a tabletop.

For other gesture types, we refer to Villarreal *et al.*'s [45,46] systematic literature reviews of GESs, although user-defined gestures for personal MDEs/MDDs have been little examined compared to cross-device input. To bridge this gap in both scientific understanding and design knowledge, we report results from two GESs, conducted to collect and analyze user-defined gestures for a 3-way display and two application domains.

3 STUDY #1: 3-WAY DISPLAYS AS REMOTE CONTROL PANELS

One popular application of gesture input is controlling remote devices in smart Internet-of-Things (IoT) environments, where input is performed either in mid-air or through a personal, mobile, or wearable device [11,18,43]. To evaluate 3-way displays as generic remote control panels for smart environments, we conducted a GES following the original method [50] and using a set of referents representative of IoT interactions in such spaces. We justify this choice for the following reasons: IoT actions are familiar for most people [43], they are frequently used in GES [45], they range from 0 to 3 dimensions, therefore enabling the participant to express preferences in a wide spectrum of possibilities.

3.1 Study

3.1.1 Participants. Twenty-four volunteers (14 females and 10 males), aged between 19 and 60 years old (M=31.8, SD=12.1, Mdn=26), were recruited via contact lists in different organizations. Their occupations included secretary, clerk, psychologist, physiotherapist, and students in law, communication, economics, sports, and management. All participants reported frequent use of computers and smartphones, no dexterity impairments, and had normal or corrected-to-normal vision. None had used a 3-way display before our study. One participant was left-handed.

3.1.2 Referents. We used a set of 19 referents representative of frequent IoT interactions performed in smart home environments, adopted from previous GEs [11,36,43,44,47]: turn TV on, turn TV off, turn alarm on, turn alarm off, turn heating on, turn heating off, turn lights on, turn lights off, turn air conditioning on, turn air conditioning off, start player, volume up, volume down, answer phone call, end phone call, go to next item in a list, go to previous item in a list, dim lights, and brighten lights (first column in Fig. 4).

3.1.3 Apparatus. The study took place in a quiet room, where information about the referents was available to the participants in visual form on a standard computer monitor. According to the principle of visual priming [20], we created visual representations of the referents as before/after states. Each representation reproduced a simplified view of our 3-way display device and highlighted the effect of a specific referent. For the 3-way display, we used a FlexLite 14" configuration composed of a Sony Vaio laptop with two expandable lateral 14" touchscreens (Fig. 1). The two panels are attached to the primary display using two hinges and are connected and powered via USB 3.0/USB-C. Each display has 1920×1080 resolution with 60 Hz frequency. Participants' gestures were video recorded. Participants were instructed to use on-screen gestures (no mid-air gestures).

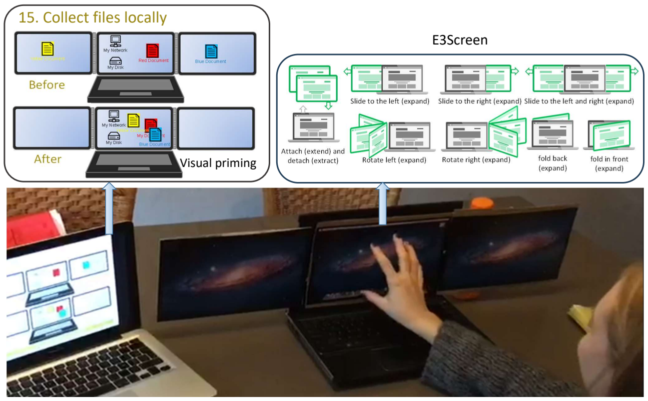


Figure 3: Setup of the study: a FlexLite 14" configuration as a main station, a second laptop showing the referents.

3.1.4 Procedure. We used a within-subjects design where participants were elicited for only one gesture per referent. After signing a consent form, participants completed a sociodemographic questionnaire, providing information about their age, gender, and handedness, as well as their use of computer technology on a 7-point Likert scale with items ranging from 1 (strongly disagree) to 7 (strongly agree). Subsequently, the participants watched an introductory video that demonstrated the features of our 3-way display device. Each session implemented the GES original protocol [50] (Fig. 3): participants were sitting in front of a 3-way display with a site laptop presenting randomly 19 *referents*, i.e., IoT operations on the E3Screen, for which they proposed suitable gestures to execute on the 3-way display those operations, i.e., gestures that fit referents well, are easy to produce and to remember. Participants were instructed to remain as natural as possible. The order of the referents was randomized per participant. The dependent variable was the agreement rate (AR) [41], a measure of the level of consensus between the proposed gestures that computes values in the $[0, 1]$ interval. We computed ARs with AGATe [41] and interpreted their magnitudes according to the thresholds of .100, .300, and .500, corresponding to low, medium, and high agreement levels.

3.2 Results

We report results on a consensus gesture set derived based on descriptive labeling [32] for each of the elicited gestures, according to Nielsen *et al.*'s [23] procedure, followed by an assignment of gestures to signs, according to the codebook model [42] of agreement analysis. We collected a total of 456 gestures corresponding to our 24 participants and set of 19 referents, which we clustered into classes of equivalence according to the gesture articulation characteristics, such as tap variations (e.g., tap *vs.* double tap) and type of stroke gestures (e.g., directional swipes or symbols), but also according to the location of the display on which the gestures were performed (left, right, or center display). Figure 4 shows the results obtained for our set of referents, in decreasing order of their corresponding AR values, together with a set of consensus gestures. Overall, ARs ranged from a maximum of .308 for "start player" to a minimum of .050 for "turn heating off," with an average level of agreement of just .140 (SD=.07, Mdn=.094), representing medium to low agreement according to our interpretation rules for ARs. Overall, we found that a percentage of 42%(=8/19) of all the referents from our study received medium agreement levels, whereas 53%(=10/19) resulted in low agreement. These findings are similar to those of previous gesture elicitation studies conducted with similar referents, e.g., Gheran *et al.* [11] reported a mean agreement rate of .112 for smart ring gestures to control devices in a smart home and Zaiţi *et al.* [54] reported .158 for mid-air gestures acquired with the Leap Motion controller for TV control. We also found that taps and stroke-gesture input, with several variations, were the most agreed gestures. Tap gestures were represented by single taps, double taps, and long/timed taps, the latter involving keeping the finger in contact with the screen for a longer period of time compared to a tap. Examples include a double tap on the right display to "turn lights on" and a timed tap to "turn the air conditioning on." The stroke gestures from the consensus set comprise of directional swipes, majoritarily performed on the center display. Although gestures were performed on all three displays, most of the articulations addressed the center and right displays. Our participants considered the center as a reference location in the personal MDD represented by the 3-way display device, while the right display was convenient to reach with the right hand since 96% of the participants in our study were right-handed.

4 STUDY #2: CONTENT MANIPULATION ON 3-WAY DISPLAYS

Touchscreen devices enable direct content manipulation through natural touch gesture input [48]. In this context, we expect a 3-way display device to enable distinctive opportunities for touch input compared to a conventional 1-way display. To understand users' preferences for gestures involving interactions with content presented on a 3-way display, we conducted another end-user gesture elicitation study. In this study, we operated with the distinction between interacting with content *within* the same display and *across* the different displays of our 3-way display device.

4.1 Study

4.1.1 Participants. Twenty-four volunteers (14 females and 10 males), aged between 12 and 68 years (M=27.9, SD=13.3, Mdn=23) were recruited following a procedure identical to that used in our first study. All of the participants reported frequent use of computers and smartphones and no dexterity impairments. None of them participated in the first study. Additionally, none of the participants had used a 3-way display prior to our study.

4.1.2 Referents. Two weeks before the GES, we asked participants to provide a list of file operations that they were using the most on their 1-way display devices. By analyzing the responses, we identified twelve tasks—"open file" (16%), "close file" (19%), "duplicate file" (10%), "delete file" (2%), "move file" (7%), "dock file" (2%), "broadcast file" (12%), "collect files" (4%), "switch files" (5%), "permute files" (2%), "merge files" (13%), and "split files" (8%),—which we used as our set of referents. Additionally, since we focused on both *within* and *across* display interactions, we considered each task performed within the limits of a single display, but also between the different displays of the 3-way display device. These combinations led to a set of twenty-four referents in our GES.

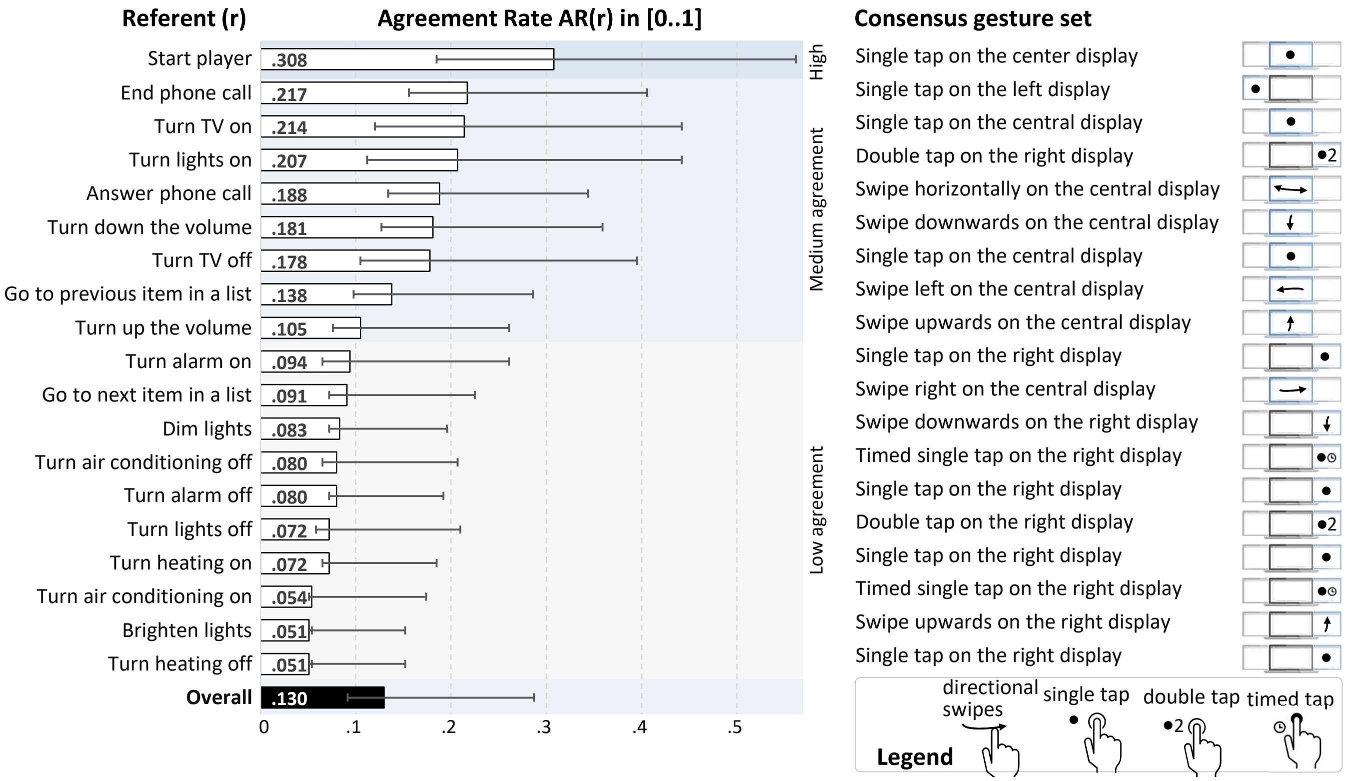


Figure 4: Consensus gesture set for a 3-way display device used as a remote control in a smart home. Error bars show 95% CIs.

4.1.3 Procedure. The apparatus and procedure were identical to the first study except for the presentation of the referents. While the referents were still randomized per participant, doubled referents, e.g., "duplicate file within" and "duplicate file across," were presented consecutively. Additionally to AR, we measured (i) Thinking-Time, defined as the time elapsed between the moment the referent was presented and the moment when the participants proposed the gesture to effect the referent, measured in seconds with a stopwatch, and (ii) Goodness-of-Fit, a rating from 1 (low) to 10 (high), expressing to what extent the participants considered their gestures appropriate to effect the referents [11].

4.2 Results

We report results on a consensus gesture set derived using the same gesture-coding procedure as in the previous study. In total, we collected 576 gestures corresponding to 24 participants and our set of 24 referents. Figure 5 shows the referents in decreasing order of their AR values, together with a consensus gesture set. Overall, ARs ranged from a maximum value of .388 for "move file across" to a minimum of .080 for "open file across," with a mean level of agreement of just .166 (SD=.148, Mdn=.141). The majority of the referents (75%=18/24) resulted in medium levels of agreement. A few exceptions aside, most of the referents received AR values similar in magnitude to those reported by previous GESs [45,46], including our first study presented in Section 3; also see [41] for a summary of AR values compiled from eighteen studies, for which the smallest level of agreement was .108 for gesture commands elicited for a MDE [32]. Unlike in our first study, however, the gesture types that constituted the consensus set were majoritarily represented by stroke gestures, such as geometrical shapes, letters, and symbols. Examples include letters "L" and "V" for merging and switching files within and across displays, and drawing a curly line symbol to close a file presented within the same display; see Figure 5 for more examples. The additional measures employed in this study led to other insights. For example, Thinking-Time measurements revealed between 8 and 20 seconds for participants to propose gestures in response to the various referents in our set, with an average of 13.4 seconds (SD=2.8, Mdn=12.9). These results indicate more time needed to propose a gesture on the 3-way display compared, for example, to mid-air gestures performed with smart rings (4.6 s) [11] and less time compared to proposing mid-air gestures for lean back interaction with the TV (20.5 s) [54]. Goodness-of-Fit measurements revealed overall high values on the 1 (low) to 10 (high) scale with a mean of 7.0 (SD=0.6, Mdn=8.3), varying between a minimum of 6.2 for "switch files within" and a maximum of 8.0 for "move file within"; see Figure 5. Furthermore, we found that most of the elicited gestures were articulated using one finger only (56% in the within and 49% in the across condition), followed by gestures performed with two fingers (28% and 30%). We also found that unistroke gestures dominated the articulations with 73% and 64%, respectively, followed by 2-stroke gestures with 20% and 27%. These findings confirm end users' preferences for simple, one-finger and unistroke touch gestures for 3-way display devices, similar to the preferences identified for other contexts of use [45,46].

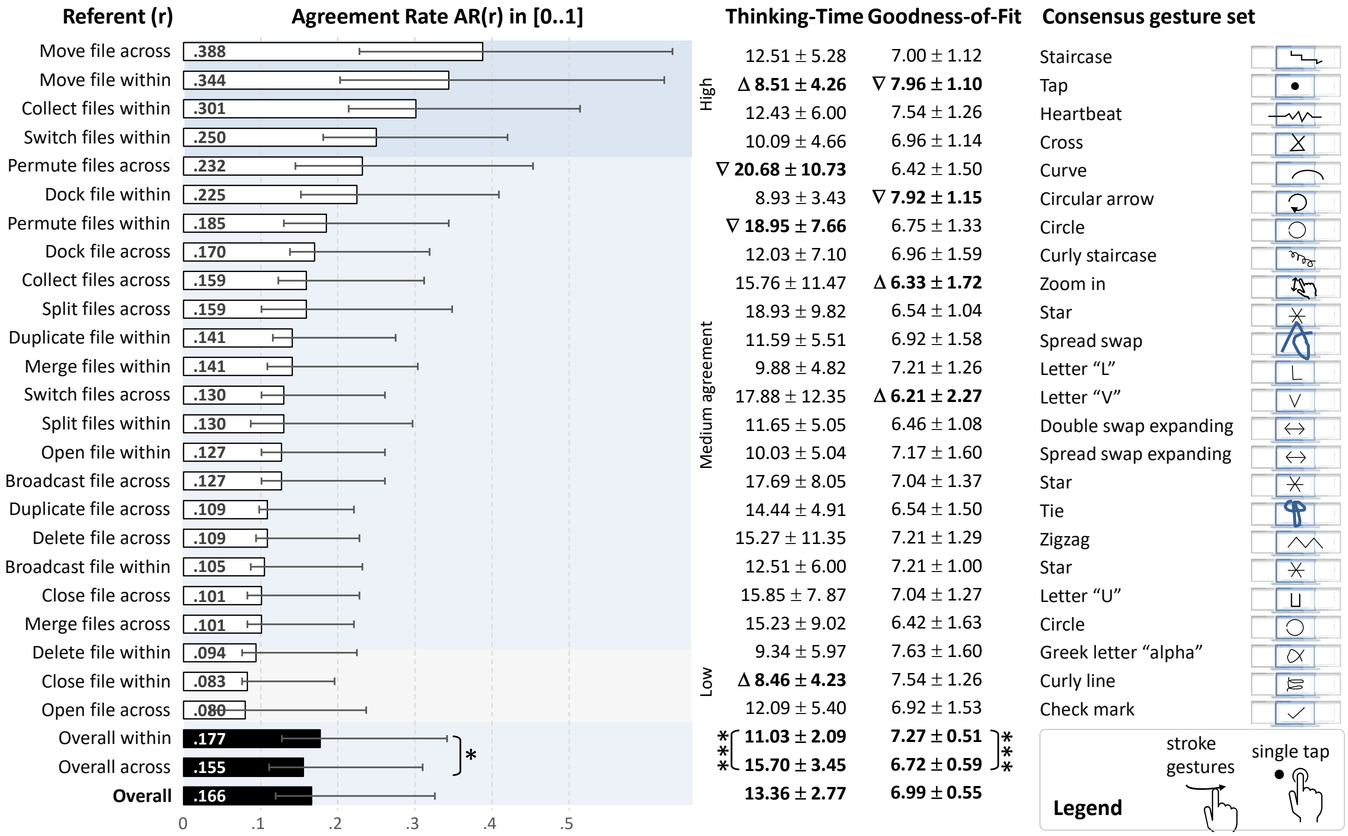


Figure 5: Consensus gesture set for content manipulation on a 3-way display device. Notes: error bars show 95% CIs; the "within" term at the end of a referent name, e.g., "merge files within," denotes an interaction taking place within the physical boundaries of the same display, while the term "across" denotes an interaction spanning across multiple displays.

5 DISCUSSION AND FUTURE WORK

Based on the findings from our two end-user gesture elicitation studies, we compiled a visual representation of typical gesture input for 3-way display devices; see Figure 6. This representation includes a total number of twenty gesture-based interactions, distributed across the three displays. For example, type-1 gestures are single-finger taps performed on any of the displays, type-2 gestures are single-finger multi-taps on the central display, while type-3 gestures include multi-taps on either lateral display. Cross-display gestures include type-11 consisting of horizontal swipes from the central display to the lateral ones, while type-12 includes single-finger swipes in the opposite direction. Starting from this basic set of gesture types, we recommend future explorations for other application domains than those considered in our two studies. To support such future work, including for personal MDDs with different numbers of displays, we also propose a taxonomy of MDE/MDD gestures, adapted from Brudy *et al.*'s [4] cross-device taxonomy; see Figure 2, left. Based on Brudy *et al.*'s possible types of cross-device systems and interactions (Figure 2a to Figure 2e), we position in this taxonomy our 3-way display gesture set as well as MDD/MDE gestures from prior work. Regarding the latter, please revisit Section 2, where we structured our discussion of related work according to the (a) to (e) categories portrayed in Figure 2. Furthermore, interesting future work opportunities are enabled by the gesture types in this space used in combination with the distinctive qualities of 3-way display form factors. For instance, the touchscreen gestures identified in our elicitation studies could be combined with manipulative gestures acting on the configurable lateral sides of a 3-way display [39] for various application domains, number of users, user categories, and contexts of use. To this end, Figure 2, right shows several opportunities for 3-way displays to be customized in terms of non-flat displays.

6 CONCLUSION

We reported findings about using 3-way display devices in two specific application scenarios: remote control of devices in a smart home environment, where the 3-way display acts as a control panel, and manipulating content on the 3-way display by leveraging its multi-display form factor for within and cross-display input. Our results revealed end users' preferences for tap gestures in the former application scenario and stroke-gesture input in the latter as well as preferences for the specific displays where gestures were articulated based on display centrality and users' handedness. To extend our preliminary findings with new discoveries, we recommend future exploration of gesture input for personal MDDs in other application domains, including augmented reality [24], in-vehicle interaction [3], and specific living environments [12].

ACKNOWLEDGMENTS

The authors thank the participants of the gesture elicitation studies reported in the paper for their contribution. Jean Vanderdonckt is supported by the EU EIC Pathfinder-Awareness Inside challenge "Symbiotik" project (1 Oct. 2022-31 Dec. 2026) under Grant no. 101071147. Radu-Daniel Vatavu acknowledges support from the NetZeRoCities Competence Center, funded by European Union - NextGenerationEU and Romanian Government, under the National Recovery and Resilience Plan for Romania, contract no. 760007/ 30.12.2022 with the Romanian Ministry of Research, Innovation, and Digitalization through the specific research project P3, Smart, and sustainable buildings. Arthur Sluÿters is funded by the "Fonds de la Recherche Scientifique - FNRS" under Grants n°40001931 and n°40011629.

REFERENCES

[1] Caroline Appert and Shumin Zhai. 2009. Using Strokes as Command Shortcuts: Cognitive Benefits and Toolkit Support. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. ACM, New York, NY, USA, 2289–2298. https://doi.org/10.1145/1518701.1519052

[2] Nathalie Aquino, Jean Vanderdonckt, Nelly Condori-Fernández, Óscar Dieste, and Óscar Pastor. 2010. Usability Evaluation of Multi-Device/Platform User Interfaces Generated by Model-Driven Engineering. In *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement* (Bolzano-Bozen, Italy) *(ESEM '10)*. ACM, New York, NY, USA, Article 30, 10 pages. https://doi.org/10.1145/1852786.1852826

[3] Laura Bianca Bilius and Radu-Daniel Vatavu. 2021. A Multistudy Investigation of Drivers and Passengers' Gesture and Voice Input Preferences for In-Vehicle Interactions. *Journal of Intelligent Transportation Systems* 25, 2 (2021), 197–220. https://doi.org/10.1080/15472450.2020.1846127

[4] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. 2019. Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Glasgow, Scotland, UK) *(CHI '19)*. ACM, New York, NY, USA, 1–28. https://doi.org/10.1145/3290605.3300792

[5] Mauricio Cirelli and Ricardo Nakamura. 2014. A Survey on Multi-touch Gesture Recognition and Multi-touch Frameworks. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces* (Dresden, Germany) *(ITS '14)*. Association for Computing Machinery, New York, NY, USA, 35–44. https://doi.org/10.1145/2669485.2669509

[6] Giuseppe Desolda, Carmelo Ardito, Hans-Christian Jetter, and Rosa Lanzilotti. 2019. Exploring spatially-aware cross-device interaction techniques for mobile collaborative sensemaking. *International Journal of Human-Computer Studies* 122 (2019), 1–20. https://doi.org/10.1016/j.ijhcs.2018.08.006

[7] Linda Di Geronimo, Marica Bertarini, Julia Badertscher, Maria Husmann, and Moira C. Norrie. 2017. Exploiting Mid-Air Gestures to Share Data among Devices. In *Proceedings of the ACM International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) *(MobileHCI '17)*. ACM, New York, NY, USA, Article 35, 11 pages. https://doi.org/10.1145/3098279.3098530

[8] Tanja Döring, Alireza Sahami Shirazi, and Albrecht Schmidt. 2010. Exploring Gesture-Based Interaction Techniques in Multi-Display Environments with Mobile Phones and a Multi-Touch Table. In *Proceedings of the ACM Conference on Advanced Visual Interfaces* (Rome, Italy) *(AVI '10)*. ACM, New York, NY, USA, 419. https://doi.org/10.1145/1842993.1843097

[9] Ana Belén Erazo and Jorge Luis Pérez Medina. 2020. Algorithmic Efficiency of Stroke Gesture Recognizers: a Comparative Analysis. *International Journal on Advanced Science, Engineering and Information Technology* 10, 2 (Mar. 2020), 438–446. https://doi.org/10.18517/ijaseit.10.2.10807

[10] Fernando Garcia-Sanjuan, Javier Jaen, and Vicente Nacher. 2016. Toward a General Conceptualization of Multi-Display Environments. *Frontiers in ICT* 3 (2016), 15 pages. https://doi.org/10.3389/fict.2016.00020 Sec. Human-Media Interaction.

[11] Bogdan-Florin Gheran, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2018. Gestures for Smart Rings: Empirical Results, Insights, and Design Implications. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 623–635. https://doi.org/10.1145/3196709.3196741

[12] Zhifan He, Ruifo Zhang, Zheng Liu, and Zhengyu Tan. 2020. A User-Defined Gesture Set for Natural Interaction in a Smart Kitchen Environment. In *Proceedings of the 13th International Symposium on Computational Intelligence and Design (ISCID '20)*. IEEE, USA, 122–125. https://doi.org/10.1109/ISCID51228.2020.00034

[13] Ken Hinckley. 2003. Synchronous Gestures for Multiple Persons and Computers. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST '03)*. ACM, New York, NY, USA, 149–158. https://doi.org/10.1145/964696.964713

[14] Ken Hinckley, Gonzalo Ramos, Francois Guimbretiere, Patrick Baudisch, and Marc Smith. 2004. Stitching: Pen Gestures That Span Multiple Displays. In *Proceedings of the ACM Working Conference on Advanced Visual Interfaces* (Gallipoli, Italy) *(AVI '04)*. ACM, New York, NY, USA, 23–31. https://doi.org/10.1145/989863.989866

[15] Christian Kray, Daniel Nesbitt, John Dawson, and Michael Rohs. 2010. User-Defined Gestures for Connecting Mobile Phones, Public Displays, and Tabletops. In *Proceedings of the ACM International Conference on Human Computer Interaction with Mobile Devices and Services* (Lisbon, Portugal) *(MobileHCI '10)*. ACM, New York, NY, USA, 239–248. https://doi.org/10.1145/1851600.1851640

[16] Ekaterina Kurdyukova, Matthias Redlin, and Elisabeth André. 2012. Studying User-Defined IPad Gestures for Interaction in Multi-Display Environment. In *Proceedings of the ACM International Conference on Intelligent User Interfaces* (Lisbon, Portugal) *(IUI '12)*. ACM, New York, NY, USA, 93–96. https://doi.org/10.1145/2166966.2166984

[17] Hyunglae Lee, Heeseok Jeong, Joongho Lee, Ki-Won Yeom, and Ji-Hyung Park. 2009. Gesture-Based Interface for Connection and Control of Multi-device in a Tabletop Display Environment. In *Proceedings of the 13th International Conference, HCI International 2009, San Diego, CA, USA, July 19-24, 2009, Human-Computer Interaction. Novel Interaction Methods and Techniques (Lecture Notes in Computer Science, Vol. 5611)*, Julie A. Jacko (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 216–225. https://doi.org/10.1007/978-3-642-02577-8_24

[18] Yihua Lou, Wenjun Wu, Radu-Daniel Vatavu, and Wei-Tek Tsai. 2017. Personalized Gesture Interactions for Cyber-Physical Smart-Home Environments. *Science China Information Sciences* 60, 7 (2017), 072104:1–15. https://doi.org/10.1007/s11432-015-1014-7

[19] Nathan Magrofuoco, Paolo Roselli, and Jean Vanderdonckt. 2021. Two-Dimensional Stroke Gesture Recognition: A Survey. *ACM Computing Survey* 54, 7, Article 155 (jul 2021), 36 pages. https://doi.org/10.1145/3465400

[20] Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, m. c. schraefel, and Jacob O. Wobbrock. 2014. Reducing Legacy Bias in Gesture Elicitation Studies. *Interactions* 21, 3 (2014), 40–45. https://doi.org/10.1145/2591689

[21] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding Users' Preferences for Surface Gestures. In *Proceedings of Graphics Interface (GI '10)*. Canadian Information Processing Society, CAN, 261–268. https://dl.acm.org/doi/10.5555/1839214.1839260

[22] Miguel A. Nacenta, Carl Gutwin, Dzmitry Aliakseyeu, and Sriram Subramanian. 2009. There and Back Again: Cross-Display Object Movement in Multi-Display Environments. *Human–Computer Interaction* 24, 1-2 (2009), 170–229. https://doi.org/10.1080/07370020902819882 arXiv:https://doi.org/10.1080/07370020902819882

[23] Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. 2004. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In *Gesture-Based Communication in Human-Computer Interaction* (Genova, Italy) *(GW '03)*, Antonio Camurri and Gualtiero Volpe (Eds.). Springer, Berlin, 409–420. https://doi.org/10.1007/978-3-540-24598-8_38

[24] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-Defined Gestures for Augmented Reality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 955–960. https://doi.org/10.1145/2468356.2468527

[25] Benjamin Poppinga, Alireza Sahami Shirazi, Niels Henze, Wilko Heuten, and Susanne Boll. 2014. Understanding Shortcut Gestures on Mobile Touch Devices. In *Proceedings of the 16th ACM International Conference on Human-Computer Interaction with Mobile Devices & Services* (Toronto, ON, Canada) *(MobileHCI '14)*. ACM, New York, NY, USA, 173–182. https://doi.org/10.1145/2628363.2628378

[26] Francesca Pulina and Fabio Paternò. 2019. Supporting cross-device interactions with gestures between personal and public devices. In *Proceedings of the 18th ACM Conference on Mobile and Ubiquitous Multimedia (MUM '19)*. ACM, New York, NY, USA, 50:1–50:5. https://doi.org/10.1145/3365610.3368419

[27] Roman Rädle, Hans-Christian Jetter, Mario Schreiner, Zhihao Lu, Harald Reiterer, and Yvonne Rogers. 2015. Spatially-Aware or Spatially-Agnostic? Elicitation and Evaluation of User-Defined Cross-Device Interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 3913–3922. https://doi.org/10.1145/2702123.2702287

[28] Gonzalo Ramos, Kenneth Hinckley, Andy Wilson, and Raman Sarin. 2009. Synchronous Gestures in Multi-Display Environments. *Human–Computer Interaction* 24, 1-2 (2009), 117–169. https://doi.org/10.1080/07370020902739288 arXiv:https://doi.org/10.1080/07370020902739288

[29] Jun Rekimoto. 1997. Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology (UIST '97)*. ACM, New York, NY, USA, 31–39. https://doi.org/10.1145/263407.263505

[30] Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-Defined Motion Gestures for Mobile Interaction. In *Proceedings of the ACM Conference on Human Factors*

in Computing Systems (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 197–206. https://doi.org/10.1145/1978942.1978971

[31] Ryan Schmidt, Eric Penner, and Sheelagh Carpendale. 2004. Reconfigurable Displays. In *Proceedings of the UbiComp'04 Workshop on Ubiquitous Display Environments (UbiDisplay '04)*. ACM, New York, NY, USA, 6 pages. https://innovis.cpsc.ucalgary.ca/innovis/uploads/Publications/Publications/ubidisplay04-schmidt-carpendale.pdf

[32] Teddy Seyed, Chris Burns, Mario Costa Sousa, Frank Maurer, and Anthony Tang. 2012. Eliciting Usable Gestures for Multi-Display Environments. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. ACM, New York, NY, USA, 41–50. https://doi.org/10.1145/2396636.2396643

[33] Teddy Seyed and Frank Maurer. 2018. *Multisurface Environments*. John Wiley & Sons, Ltd, West Sussex, UK, Chapter 22, 515–538. https://doi.org/10.1002/9781118976005.ch22

[34] Vivian Shen and Chris Harrison. 2022. Pull Gestures with Coordinated Graphics on Dual-Screen Devices. In *Proceedings of the 2022 International Conference on Multimodal Interaction* (Bengaluru, India) *(ICMI '22)*. ACM, New York, NY, USA, 270–277. https://doi.org/10.1145/3536221.3556620

[35] Adalberto L. Simeone, Julian Seifert, Dominik Schmidt, Paul Holleis, Enrico Rukzio, and Hans Gellersen. 2013. A Cross-Device Drag-and-Drop Technique. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia* (Luleå, Sweden) *(MUM '13)*. ACM, New York, NY, USA, Article 10, 4 pages. https://doi.org/10.1145/2541831.2541848

[36] Arthur Sluÿters, Quentin Sellier, Jean Vanderdonckt, Vik Parthiban, and Pattie Maes. 2023. Consistent, Continuous, and Customizable Mid-Air Gesture Interaction for Browsing Multimedia Objects on Large Displays. *International Journal of Human–Computer Interaction* 39, 12 (2023), 2492–2523. https://doi.org/10.1080/10447318.2022.2078464 arXiv:https://doi.org/10.1080/10447318.2022.2078464

[37] Nikita Soni, Schuyler Gleaves, Hannah Neff, Sarah Morrison-Smith, Shaghayegh Esmaeili, Ian Mayne, Sayli Bapat, Carrie Schuman, Kathryn A. Stofer, and Lisa Anthony. 2019. Do User-Defined Gestures for Flatscreens Generalize to Interactive Spherical Displays for Adults and Children?. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays* (Palermo, Italy) *(PerDis '19)*. ACM, New York, NY, USA, Article 24, 7 pages. https://doi.org/10.1145/3321335.3324941

[38] Kênia Sousa, Hildeberto Mendonça, Jean Vanderdonckt, Els Rogier, and Joannes Vandermeulen. 2008. User Interface Derivation from Business Processes: A Model-Driven Approach for Organizational Engineering. In *Proceedings of the ACM Symposium on Applied Computing* (Fortaleza, Ceara, Brazil) *(SAC '08)*. ACM, New York, NY, USA, 553–560. https://doi.org/10.1145/1363686.1363821

[39] Jean Vanderdonckt and Radu-Daniel Vatavu. 2021. Extensible, Extendable, Expandable, Extractable: The 4E Design Approach for Reconfigurable Displays. *International Journal of Human-Computer Interaction* 37, 18 (2021), 1720–1736. https://doi.org/10.1080/10447318.2021.1908666

[40] Radu-Daniel Vatavu, Annette Mossel, and Christian Schönauer. 2016. Digital Vibrons: Understanding Users' Perceptions of Interacting with Invisible, Zero-Weight Matter. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, New York, NY, USA, 217–226. https://doi.org/10.1145/2935334.2935364

[41] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 1325–1334. https://doi.org/10.1145/2702123.2702223

[42] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2022. Clarifying Agreement Calculations and Analysis for End-User Elicitation Studies. *ACM Trans. Comput.-Hum. Interact.* 29, 1, Article 5 (jan 2022), 70 pages. https://doi.org/10.1145/3476101

[43] Radu-Daniel Vatavu and Ionut-Alexandru Zaiti. 2014. Leap Gestures for TV: Insights from an Elicitation Study. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (Newcastle Upon Tyne, United Kingdom) *(TVX '14)*. ACM, New York, NY, USA, 131–138. https://doi.org/10.1145/2602299.2602316

[44] Santiago Villarreal-Narvaez, Alexandru-Ionuţ Şiean, Arthur Sluÿters, Radu-Daniel Vatavu, and Jean Vanderdonckt. 2022. Informing Future Gesture Elicitation Studies for Interactive Applications That Use Radar Sensing. In *Proceedings of the ACM International Conference on Advanced Visual Interfaces* (Frascati, Rome, Italy) *(AVI '22)*. ACM, New York, NY, USA, Article 50, 3 pages. https://doi.org/10.1145/3531073.3534475

[45] Santiago Villarreal-Narvaez, Arthur Sluÿters, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2024. Brave New GES World: A Systematic Literature Review of Gestures and Referents in Gesture Elicitation Studies. *ACM Comput. Surv.* 56, 5, Article 128 (jan 2024), 55 pages. https://doi.org/10.1145/3636458

[46] Santiago Villarreal-Narvaez, Jean Vanderdonckt, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2020. A Systematic Review of Gesture Elicitation Studies: What Can We Learn from 216 Studies?. In *Proceedings of the ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. ACM, New York, NY, USA, 855–872. https://doi.org/10.1145/3357236.3395511

[47] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. 2022. *'Address and command': Two-handed mid-air interactions with multiple home devices. International Journal of Human-Computer Studies* 159 (2022), 102755. https://doi.org/10.1016/j.ijhcs.2021.102755

[48] Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. https://dl.acm.org/doi/10.5555/1995309

[49] Andrew D. Wilson and Hrvoje Benko. 2010. Combining Multiple Depth Cameras and Projectors for Interactions on, above and between Surfaces. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 273–282. https://doi.org/10.1145/1866029.1866073

[50] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-Defined Gestures for Surface Computing. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[51] Haijun Xia, Michael Glueck, Michelle Annett, Michael Wang, and Daniel Wigdor. 2022. Iteratively Designing Gesture Vocabularies: A Survey and Analysis of Best Practices in the HCI Literature. *ACM Trans. Comput. Hum. Interact.* 29, 4 (2022), 37:1–37:54. https://doi.org/10.1145/3503537

[52] Zhican Yang, Chun Yu, Xin Chen, Jingjia Luo, and Yuanchun Shi. 2022. Investigating user-defined flipping gestures for dual-display phones. *International Journal of Human-Computer Studies* 163 (2022), 102800. https://doi.org/10.1016/j.ijhcs.2022.102800

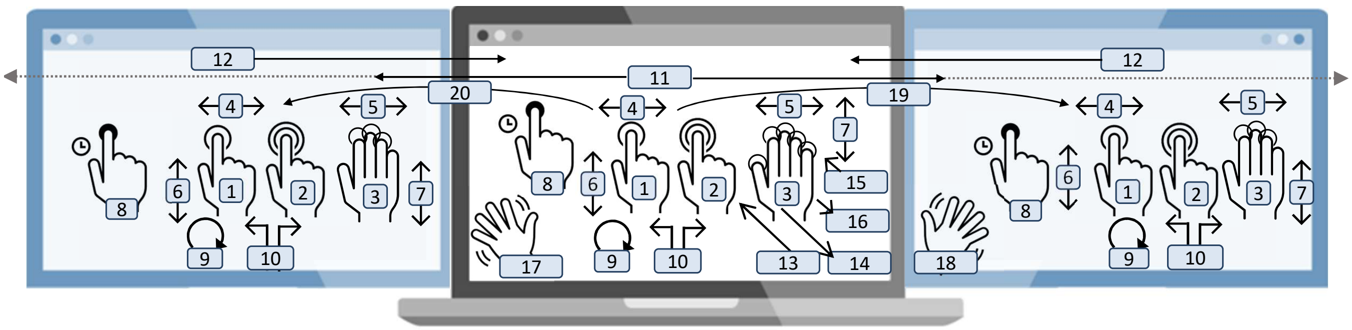
[53] Yen-Ting Yeh, Fabrice Matulic, and Daniel Vogel. 2023. Phone Sleight of Hand: Finger-Based Dexterous Gestures for Physical Interaction with Mobile Phones. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. ACM, New York, NY, USA, Article 519, 19 pages. https://doi.org/10.1145/3544548.3581121

[54] Ionut-Alexandru Zaiti, Stefan-Gheorghe Pentiuc, and Radu-Daniel Vatavu. 2015. On Free-Hand TV Control: Experimental Results on User-Elicited Gestures with Leap Motion. *Personal and Ubiquitous Computing* 19, 5-6 (2015), 821–838. https://doi.org/10.1007/s00779-015-0863-y

[55] Shumin Zhai, Per Ola Kristensson, Caroline Appert, Tue Haste Anderson, and Xiang Cao. 2012. Foundational Issues in Touch-Surface Stroke Gesture Design — An Integrative Review. *Foundations and Trends in Human–Computer Interaction* 5, 2 (2012), 97–205. https://doi.org/10.1561/1100000012

A CLASSIFICATION OF TOUCHSCREEN GESTURES FOR A 3-WAY DISPLAY

This appendix shows the classification resulting from the studies.



1 Single-finger single-tap on either display
2 Single-finger multi-tap on either display
3 Multi-finger single-tap on either display
4 Single-finger horizontal swipe on either display
5 Multi-finger horizontal swipe on either display
6 Single-finger vertical swipe on either display
7 Multi-finger vertical swipe on either display
8 Long/timed single tap on either display
9 Single-finger geometrical shape on either display
10 Single-finger symbol on either display
11 Horizontal swipe from the central to a lateral display
12 Horizontal swipe from a lateral display to the central display
13 Single-finger directional swipe from an edge to the center
14 Single-finger directional swipe from the center to an edge
15 Single-finger directional swipe from an edge
16 Single-finger directional swipe from the center
17 Multi-finger pinch in/out on the central display
18 Multi-finger pinch in/out extended from the center to a lateral display
19 Multi-finger pinch in/out from center to right
20 Multi-finger pinch in/out from center to left

Figure 6: Our classification of possible touchscreen gesture interactions for 3-way display.