Congruent and Hierarchical Gesture Set Design

Arthur Sluÿters
Université catholique de Louvain, LouRIM
Louvain-la-Neuve, Belgium
arthur.sluyters@uclouvain.be

Jean Vanderdonckt
Université catholique de Louvain, LouRIM
Louvain-la-Neuve, Belgium
jean.vanderdonckt@uclouvain.be

Paolo Roselli
Università degli Studi di Roma “Tor Vergata”
Dipartimento di Matematica
Roma, Italy
roselli@mat.uniroma2.it

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

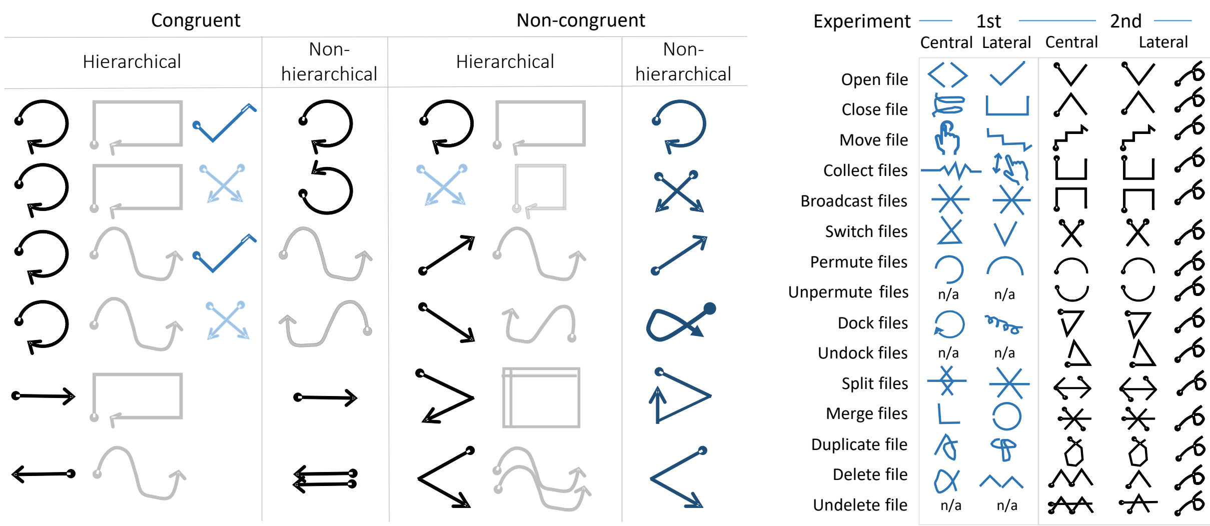

Figure 1: Non-congruent, non-hierarchical stroke gestures *vs.* congruent and hierarchical alternatives (*left*); The results of our experiments comparing non-congruent, non-hierarchical gestures and their congruent, hierarchical counterparts (*right*).

Abstract

The typical approach to gesture set design, which relies on one-to-one mappings between gestures and system functions, often presents challenges for users in terms of gesture discoverability, learnability, and memorability. In this paper, we examine the hypothesis that semantically related system functions can benefit from the use of *congruent gestures,* whereas functions structured in the form of parameterized action may be better supported by *hierarchical gestures.* We report the results of a gesture elicitation study conducted with n_1=24 participants, who proposed stroke gestures for a multi-display touchscreen to effect file-related manipulation referents either locally on a central display or remotely on a lateral display. In a follow-up study, an original mixed method combining elicitation and identification, another sample of n_2=24 participants was instructed to focus on congruent and hierarchical gestures for the same referents. Our results reveal higher agreement and an increased perceived goodness of fit between gestures and system functions in the second study.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; **Gestural input**; **User studies**.

Keywords

Congruent gestures, Hierarchical gestures, Gesture elicitation, Identification study, Mixed elicitation, Stroke gestures, Touchscreen

ACM Reference Format:
Arthur Sluÿters, Jean Vanderdonckt, Paolo Roselli, and Radu-Daniel Vatavu. 2025. Congruent and Hierarchical Gesture Set Design. In *Designing Interactive Systems Conference (DIS '25 Companion), July 5–9, 2025, Funchal, Portugal.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3715668.3736383

DIS '25 Companion, Funchal, Portugal

ACM ISBN 979-8-4007-1486-3/2025/07
https://doi.org/10.1145/3715668.3736383

1 Introduction

The number of gestures users can effectively memorize and recall depends on various human factors, such as cognitive abilities and learning capacity, as well as on the gesture-to-function mappings within the context of use of the interactive system featuring the gesture-based interface. For example, gestures that have meaningful associations with the functions they trigger, have simple forms, or replicate physical actions from the real world tend to be more memorable. In the case of touch and stroke gesture input [41], gesture sets are typically determined by the operating system (e.g., seven default gestures in Apple iOS and fourteen gestures in Microsoft Surface) or the sensing and recognition capabilities of the device (e.g., twelve gestures available when using Kinemic). These limitations can be overcome by implementing applications with custom gesture recognition solutions [13,15–17,30].

The assignment of gestures to system commands [3,24] determines key usability factors, such as discoverability [4], memorability [6,10,21], articulation performance [8], learning [3,21,22], and recall [3,6]. Although the design of a gesture set should consider contextual, cognitive, physical, and system factors [39], the typical approach, despite some exceptions [9,23], still follows a one-to-one mapping [6], where a single gesture is assigned to each command. Consequently, as the size of the command set increases, the expected benefits tend to decrease. For example, during gesture elicitation [38], participants may run out of options, revert to previously known interactions [19], or create unnecessarily complex gestures with abstract associations to commands to complete the task. This can negatively impact gesture set memorability and recall, with the problem potentially exacerbated in many-to-one mappings.

A potential solution lies in *gesture congruency* (where related system functions are assigned related gestures) and *gesture hierarchical structures* (where gestures take the form of action verb-object and action-qualifier-object). A representative example of non-congruent gestures occurs when users assign unrelated gestures to opposite, symmetric, or dichotomous commands, e.g., "draw a circle" and "draw a cross" in mid-air assigned to "turn TV on" and "turn TV off", respectively; see the fourth column in Figure 1. In contrast, a "clockwise circle" followed by "rectangle" and "cross" followed by "square" result in hierarchical gestures that can be meaningfully decomposed into constituent parts, although they still remain non-congruent themselves; see the third column in Figure 1. Moreover, the complexity of the proposed gestures sometimes exceeds the number of dimensions of the command: e.g., a unistroke and a multistroke gesture were proposed for a two-state on/off action.

As alternatives to non-congruent, non-hierarchical commands, we explore in this paper *congruent and hierarchical gestures*; see the first column in Figure 1. We believe that gestures with these characteristics preserve the intuitive quality of gesture input, have strong cognitive associations to system functions, and enhance both memorability and recall. Section 2 reviews previous work on the composition of system commands and gestures. Section 3 reports findings from a gesture elicitation study that produced a set of non-congruent, non-hierarchical gestures for specific referents involving E3Screen, a multi-display device [28,29]. Section 4 presents the findings of a mixed elicitation study, focusing on both congruent and hierarchical gestures for the same referents.

2 Related Work

Carroll [5] compared four versions of a 16-command language for human-robot interaction, based on whether the commands were *congruent* (i.e., symmetric for pairs of semantically related commands) and *hierarchical* (i.e., following a verb-object-qualifier structure). For example, "advance/retreat" and "right/left" are congruent commands, while "go/back" and "turn/left" are not. Congruent and hierarchical commands received the best subjective ratings, the highest test scores on a problem-solving task, and the fewest errors and omissions. Carroll [5] conjectured that recall would improve due to the congruent and hierarchical structure of these commands.

To transpose this design principle to gesture input, we need to examine the process of *gesture composition*, for which various approaches exist, such as chaining primitives [14], spatio-temporal combination [6], concatenation [24], contiguity [12], and assembly [37]. For example, hierarchical gestures are well-suited for smartphone control when the number of commands is large, as they offer better learnability, higher expressiveness, and greater subjective satisfaction compared to non-hierarchical variants [14]. Another example is the Augmented Letters technique [24], where commands are formed by composing unistrokes and parameters, e.g., letter "T" for turn followed by a right flick to indicate the turning direction. Delamare et al. [6] apply the same principle to 3D mid-air input to derive a vocabulary of hierarchical gestures with an improved recall rate. User-defined gestures are generally easier to remember than those created by designers or random mappings [21]. End users prefer gestures proposed by large groups [20].

There are also frameworks for designing gesture-based interfaces across various applications. For example, Yao et al. [40] proposed a layered architecture for multi-touch gestures applied to urban planning, comprising raw data, basic gesture, and application-specific layers. Similarly, Acuna et al. [1] describe a multilayered framework for two-handed gestures in 3D software development environments. [7] introduced a layered gesture recognition framework that provides device independence and extensibility for 3D gestures captured with gloves. Mo and Neumann [18] present a framework for automatically producing a gesture interface based on a simple interface description written in Gelex notation, in which each hand pose is decomposed into elements in a finger-pose alphabet. These frameworks simplify the design and implementation of gesture interfaces by providing structured approaches to their recognition and interpretation and by offering flexibility in their layered architectures for different gesture types.

Prior research suggests that gesture elicitation studies [38], a specific form of participatory design where participants propose desirable gestures corresponding to given system functions, can be an effective method for investigating the congruent and hierarchical properties of gesture input. Vogiatzidakis and Koutsabasis [37] showed that prioritizing gestures according to a command-and-address structure improved their acceptability and usage. Apart from this study and to the best of our knowledge, no research has explored the congruent and hierarchical properties in the context of touch and stroke gesture input. While knowledge in user-defined gestures has been accumulating [25–27,34–36], there is limited understanding about gestures that exhibit congruency and hierarchical properties in relation to the system functions they effect.

3 Experiment #1: Gesture Elicitation Study

We conducted a gesture elicitation study, following the method introduced by Wobbrock *et al.* [38], to understand the congruent and hierarchical characteristics of user-defined gestures. We place this study in the context and application area of interactive multi-display environments [26], and elicited touch and stroke gesture input for E3Screen [28,29], a reconfigurable touchscreen display compatible with tablets, laptops, and desktop monitors.

3.1 Participants

Twenty-four participants (fourteen female and ten male, aged from 12 to 68 years, *M*=27.9, *SD*=13.3, *Mdn*=23), were recruited for the study through a contact list in different organizations and convenience sampling. The wide age range (a span of 56 years) supported the exploratory nature of our study, and for the 12-year-old participant, we obtained the consent of their legal guardian. Participants' occupations were diverse, including secretary, administrative clerk, psychologist, physical therapist, and law, communication, economics, sports, and management students. All participants reported regular use of computers and smartphones. None had seen or used an E3Screen or other similar multi-display device before.

3.2 Referents

Prior to the experiment, we asked participants to enumerate the most frequent file-related tasks they used on a regular basis, which resulted in the following set of twelve referents: *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%). Since E3Screen consists of three displays (see details in [28,29]), participants were instructed to propose gestures either locally on the central display or remotely, on the lateral screens. We applied visual priming [20] in the form of interactive PowerPoint presentations of the 12 tasks × 2 locations (central *vs.* lateral) = 24 referents showing before/after illustrations of successfully completed system function effects.

3.3 Setup

A within-subjects design was implemented where each participant proposed one gesture per referent. The experiment took place in a quiet room. Illustrations of the referents were presented on a laptop next to E3Screen (see Figure 2). All gestures were recorded using both a screen-recording application and a video camera.

3.4 Procedure and Task

Participants signed an informed consent form, completed a socio-demographic questionnaire (age, gender, handedness, use of computer technology and digital devices), and performed the NEPSY-II motor skill test [11] consisting of pinching each finger with the thumb several times in a row. Subsequently, the participants were presented with the referents, in randomized order, for which they proposed suitable gestures to maximize the goodness of fit, ease of articulation, and ease of recall. The referents were presented in pairs according to the display location, e.g., *duplicate file locally* and *duplicate file remotely*. Each session lasted about 45 minutes. We measured the following dependent variables:

- Agreement-Rate (AR), ratio variable, expressing the agreement among participants' gestures, using AGATe [31].
- Thinking-Time (TT), a ratio variable, defined as the time (in seconds) elapsed between the moment a referent was first shown and the participant confirmed having found a suitable gesture to perform it, measured using a stopwatch.
- Goodness-of-Fit (GoF), an ordinal variable with values ranging from 1 (low) and 10 (high), indicating the extent to which participants perceived the gesture they proposed as suitable for effecting the corresponding referent.

3.5 Results

The gestures were subjected to descriptive labeling [26], following Nielsen *et al.*'s [22] procedure, and assigned to a gesture category, according to the codebook model to gesture analysis in end-user elicitation studies [32]. A total of 576 gestures were clustered into 26 equivalence classes, as follows: *pinch*, *right arrow*, *triangle*, *zigzag*, *heartbeat-like pattern*, *knot*, *check mark*, *letter "U,"* *star*, *spiral*, *letter "L,"* *curve*, *stair-like pattern*, *cross*, *pinch out*, *circle*, *double swap*, *pinch in*, *spread swap*, *tap*, *double tap*, *cloud*, *press once*, *square*, *infinity symbol*, and *letter "D."* Since the design space of possible gestures afforded by multi-screen devices is wide, participants came up with many different proposals. Figure 3, left shows AR values for each referent. Overall, they ranged from low agreement (e.g., .08 for *open file remotely*) to medium agreement (.388 for *move file remotely*) with a mean of .166 on the unit scale (*SD*=.148). Thinking-Time ranged from 8.5s for *close file locally* to 20.7s for *permute files locally* (*M*=13.5, *SD*=3.3); see Figure 3, middle. A Wilcoxon signed-rank test revealed a statistically significant difference between referents corresponding to the local (*M*=11.0) and remote (*M*=15.7) displays of E3Screen. Goodness-of-Fit ratings, averaged across participants, ranged from a minimum of 6.2 for *switch files locally* to 8.0 for *move file locally* (*M*=6.9, *SD*=1.4); see Figure 3, right. A Wilcoxon signed-rank test found a significant difference (*Z*=1.79, *p*=.036) for local (*M*=7.1, *SD*=1.4) *vs.* remote (*M*=6.6, *SD*=1.6) displays.

4 Experiment #2: Mixed Elicitation Study

4.1 Method

In our second study, we combined end-user elicitation [38] and identification [2] by presenting participants with a gesture set, while also offering them the opportunity to propose alternative gestures. We started from the consensus gesture set identified in the first study and employed the same apparatus and task, as described in Section 3. We recruited a new sample of 24 participants (seventeen male and seven female, aged from 16 to 60 years, *M*=32.01, *SD*=8.08, *Mdn*=31). The age range with a span of 44 years was smaller, but similar to the one in the first study, and the 16-year old participant had the consent of their legal guardian. None of the participants were involved in the first study. A set of 30 referents was composed based on the 24 initial ones, which were expanded with new referents corresponding to congruent and hierarchical design, e.g., *undelete file* for the *delete file* referent, *undock file* for *dock file*. The participants could select a gesture from the provided set (identification) or propose a new gesture (elicitation). The presentation order of the referents was randomized. We collected 24 (participants) × 15 (referents) × 2 (display locations) = 720 gesture proposals.

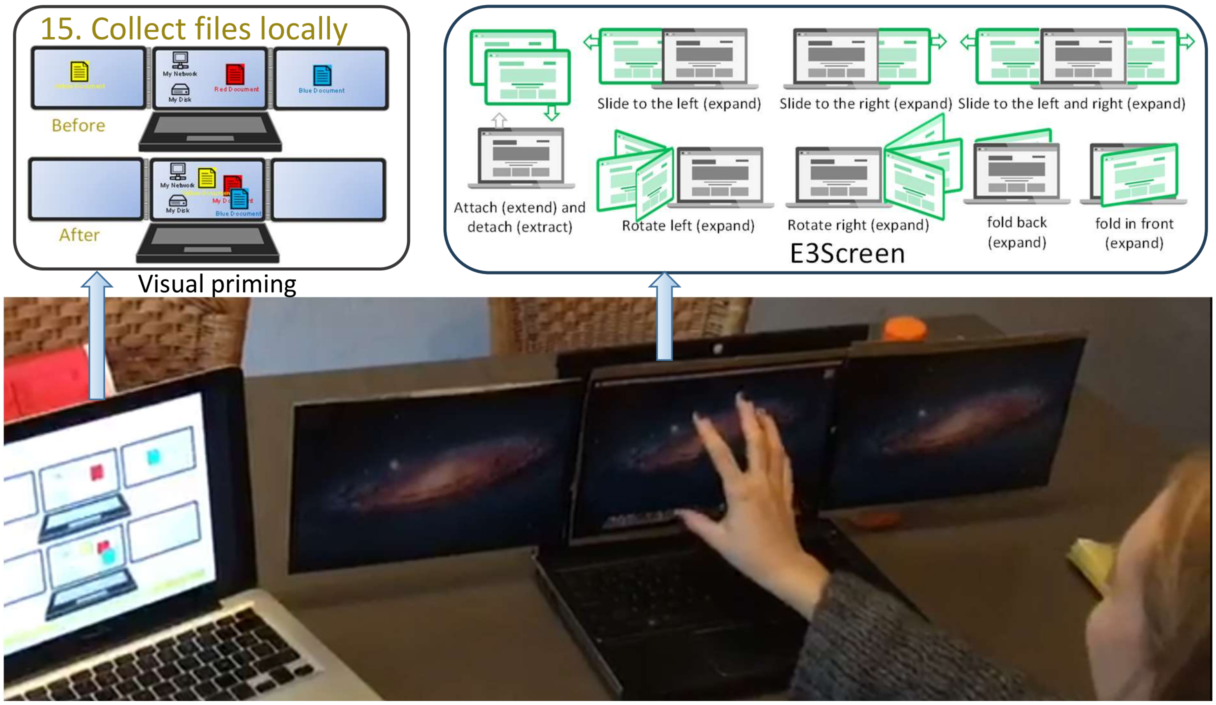

Figure 2: Setup of the design study with the E3Screen as a source for gesture input.

4.2 Results

Figure 3 shows the results obtained in the second study (light colors) *vs.* the ones in first study (dark colors) in terms of Agreement-Rate (left), Thinking-Time (middle), and Goodness-of-Fit (right). Overall, the average agreement increased from .166 to .281 from the first to the second study, representing an increase of 69%. A two-tailed Mann-Whitney test for independent samples showed that this difference was statistically significant (U=132, z=3.96, p<.001) with a large effect size (r=.54). When restricting the comparison to the common referents only, the test remained significant (U=114, z=3.58, p<.001, r=.51). The difference in the referents involving the local, central display was 59% and that between the referents involving the lateral, remote displays was 81%. These results suggest that the second study resulted in higher agreement overall when congruent and hierarchical gestures were involved.

The average Thinking-Time decreased from 13.5 s to 12.5 s, a reduction of 8%, not statistically significant (U=396, z=0.79, p=.220>.05, *n.s.*). Similarly, the decrease in Thinking-Time for local (1%) and remote (12%) displays was not statistically significant (p=.820 and p=.070, *n.s.*). These results reveal that participants spent similar amounts of thinking time on congruent, hierarchical gestures as the other sample did on non-congruent, non-hierarchical ones.

The average Goodness-of-Fit increased from 6.9 to 8.2, on the 1 (low) to 10 (high) scale, representing an improvement of 18%. A two-tailed Mann-Whitney test for independent samples showed that this difference was statistically significant (U=76, z=5.52, p<.001) with a large effect size (r=.71). When restricting the comparison to the common referents only, the test remained significant (U=50, z=4.90, p<.001, r=.70). The difference in referents for the local, central display between the two studies was 19%, which was also statistically significant (U=2.5, z score=4.25, p<.001***) with a large effect size (r=.81). The difference for the lateral, remote displays was 17%, which was also statistically significant (U=19.5, z score=3.42, p=.00023***) with a large effect size (r=.66). These results indicate that participants expressed greater satisfaction in the mappings involving gestures exhibiting congruent and hierarchical properties than with those that did not have these characteristics.

5 Application

We should prefer congruent, hierarchical gestures when dealing with a large number of commands or referents that are related to each other or connected between them, that is, when the gesture-based user interface involves many objects, functions, or categories sharing a common meaning. This is particularly the case in ambient intelligence when many functions (e.g., "Turn on", "Turn off") are similar across many objects (e.g., "lamp") or devices (e.g., "TV", "media player", "heating system"), such as in multi-device or multi-platform contexts of use. Indeed, these contexts intrinsically manipulate many devices (e.g., a smartphone, a tablet, a laptop, a desktop, a tabletop) sharing many similar functions that are equally applicable to any device. So, instead of repeating the same function on different devices through a different gesture each time, it is more beneficial to highlight the common areas to determine the gestures and to map a gesture to one common function or device. Hierarchical gestures allow end users to navigate or recall commands by breaking them into meaningful groups, reducing cognitive load. Congruent gestures, where the gesture resembles the referent (e.g., a pinch-out gesture to "zoom in") further aids memorability and

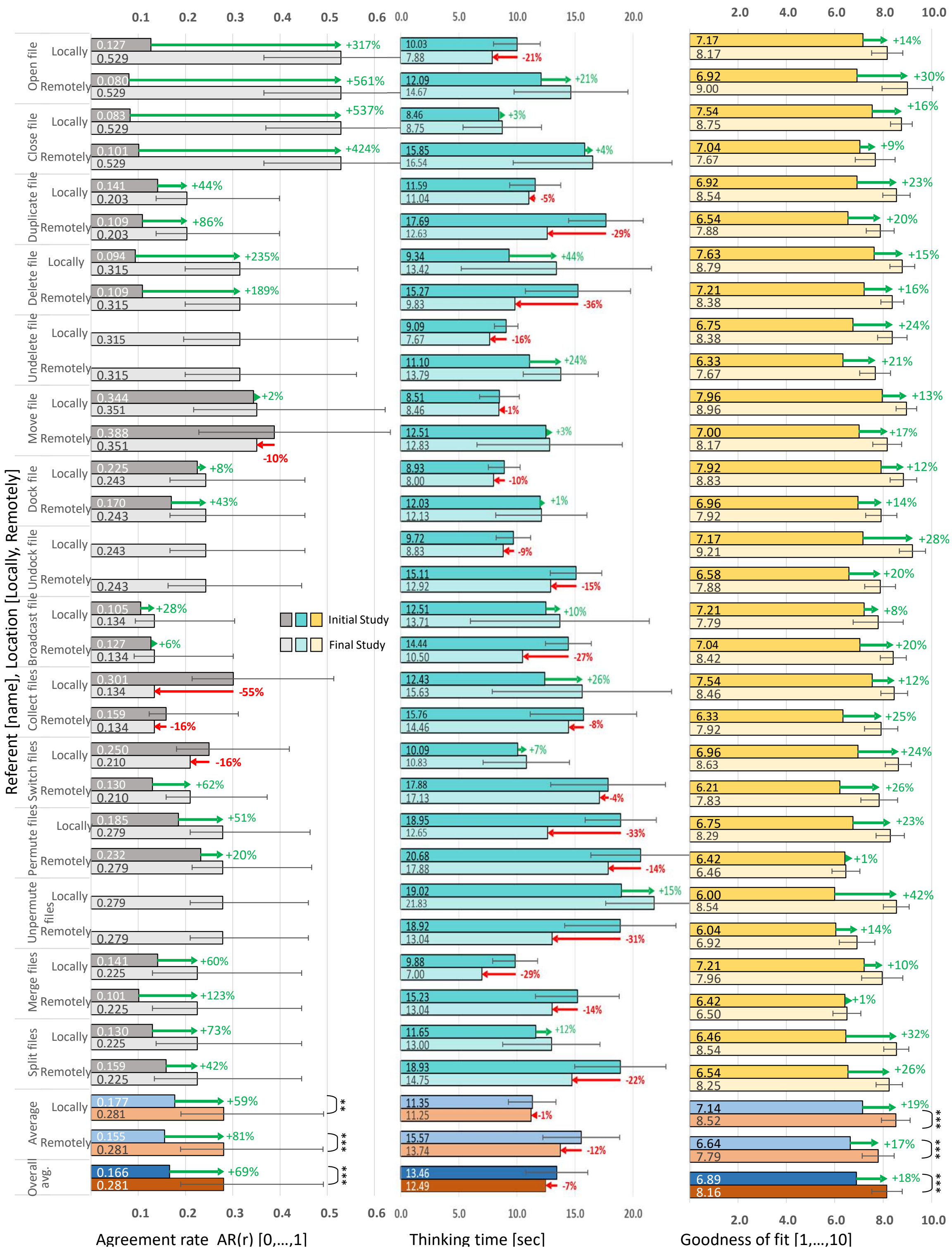

Figure 3: Comparison between the results obtained during the two gesture studies: Agreement-Rate (*left*), Thinking-Time (*middle*), and Goodness-of-Fit (*right*). Note: error bars show 95% confidence intervals.

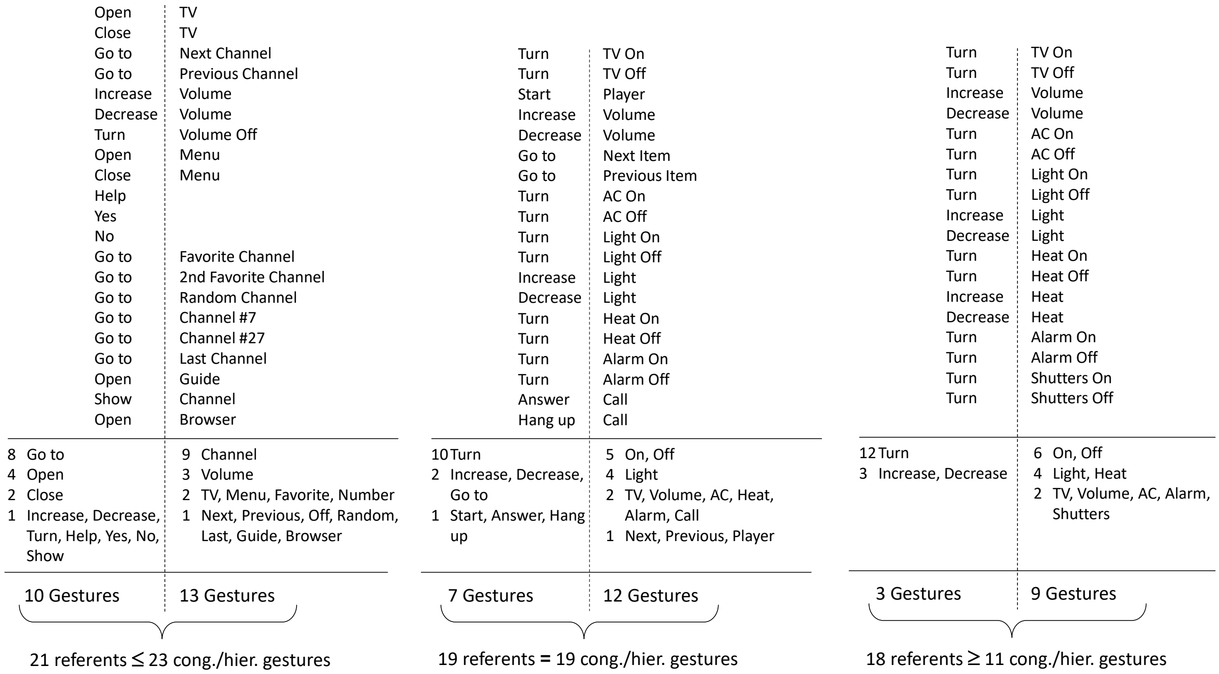

Figure 4: Three cases of application of congruent, hierarchical gestures.

intuitiveness, making it easier to learn and retain a large set of commands and their associated gestures.

Conversely, when the number of commands or referents is small, non-congruent, non-hierarchical gestures may suffice or even be preferable. Simpler gesture-based user interfaces do not require extensive categorization, and individual gestures can be efficient once learned. However, we do not know today the threshold beyond which the decision in favor of congruent, hierarchical gestures should be taken.

Figure 4 shows three application cases: when the number of referents is smaller than the number of congruent/hierarchical gestures, based on Vatavu and Zaiti [33] (left), when the number of referents is equal to the number of gestures, based on Magrofuoco et al. [15] (center), and when the number of referents is larger than the number of gestures (right).

Figure 4-left shows a case of application with 21 referents, ranging from "Open TV" to "Open Browser". An analysis of the frequency of appearance of each symbol in the vocabulary of these 21 referents reveals which are the most frequent (e.g., 9 times "Channel" and 8 times "Go to", which are good candidates), which are less frequent (e.g., 2 times "Close"), and which appear only once (e.g., "Increase", "Decrease", which are not good candidates since they are not repeated or associated). In this case, the number of individual gestures associated with each symbol of the vocabulary exceeds the initial number of referents (23 vs. 21). Conversely, Figure 4-right exhibits less symbols with higher frequency (e.g., 12 times "Turn", 6 times "on, off", which are semantically related), thus resulting in a number of individual gestures smaller than the number of referents (11 vs. 18). Applying the congruent-hierarchical gestures therefore depends on the vocabulary configuration.

6 Conclusion

We reported insights into the congruent and hierarchical characteristics of user-defined gestures based on two studies involving 48 participants. We noted a 69% increase in agreement rate, similar thinking times, and an 18% improvement in the perceived goodness-of-fit between gestures and their corresponding functions when gestures exhibited congruency and hierarchical structure. Although these findings are specific to the multi-display environment evaluated in our studies, we believe that congruent and hierarchical gestures have the potential to enhance user performance and experience across other application domains as well. We recommend future studies to replicate our findings in new contexts of use and an experimental study to determine an actionable guideline specifying when the congruent-hierarchical gestures are applicable. This could be determined by a quantitative measure based on the vocabulary tree.

Acknowledgments

The authors thank the participants of the gesture studies reported in this article. Arthur Sluÿters was funded by the Fonds de la Recherche Scientifique - FNRS under Grants no. 40001931 and no. 40011629. Jean Vanderdonckt is supported by the EU EIC Pathfinder-Awareness Inside challenge Symbiotik project (1 Oct. 2022-30 Sept. 2026) under Grant no. 101071147. Radu-Daniel Vatavu acknowledges support from the NetZeRoCities Competence Center funded by the European Union-NextGenerationEU and the 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.

References

[1] Raul A. Herrera Acuna, Christos Fidas, Vasileios Argyriou, and Sergio A. Velastin. 2012. Toward a Two-Handed Gesture-Based Visual 3D Interactive Object-Oriented Environment for Software Development. In *Proceedings of Eighth International Conference on Intelligent Environments*. 359–362. https://doi.org/10.1109/IE.2012.28

[2] Abdullah X. Ali, Meredith Ringel Morris, and Jacob O. Wobbrock. 2019. Crowdlicit: A System for Conducting Distributed End-User Elicitation and Identification Studies. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Glasgow, Scotland, UK) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 255, 12 pages. https://doi.org/10.1145/3290605.3300485

[3] 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)*. Association for Computing Machinery, New York, NY, USA, 2289–2298. https://doi.org/10.1145/1518701.1519052

[4] Francesco Cafaro, Leilah Lyons, and Alissa N. Antle. 2018. Framed Guessability: Improving the Discoverability of Gestures and Body Movements for Full-Body Interaction. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174167

[5] John M. Carroll. 1982. Learning, using and designing filenames and command paradigms. *Behaviour & Information Technology* 1, 4 (1982), 327–346. https://doi.org/10.1080/01449298208914457

[6] William Delamare, Chaklam Silpasuwanchai, Sayan Sarcar, Toshiaki Shiraki, and Xiangshi Ren. 2019. On Gesture Combination: An Exploration of a Solution to Augment Gesture Interaction. In *Proceedings of the ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 135–146. https://doi.org/10.1145/3343055.3359706

[7] J. Eisenstein, S. Ghandeharizadeh, L. Golubchik, C. Shahabi, Donghui Yan, and R. Zimmermann. 2003. Device independence and extensibility in gesture recognition. In *Prcoeedings of the IEEE International Conference on Virtual Reality* (Los Angeles, CA, USA). IEEE Press, Piscataway, NJ, USA, 207–214. https://doi.org/10.1109/VR.2003.1191141

[8] Orlando Erazo, José A. Pino, and Pedro Antunes. 2015. Estimating Production Time of Touchless Hand Drawing Gestures. In *Proceedings of 15th IFIP TC 13 International Conference on Human-Computer Interaction, INTERACT '15* (Bamberg, Germany) *(Lecture Notes in Computer Science, Vol. 9298)*, Julio Abascal, Simone D. J. Barbosa, Mirko Fetter, Tom Gross, Philippe A. Palanque, and Marco Winckler (Eds.). Springer, Cham, 552–569. https://doi.org/10.1007/978-3-319-22698-9_38

[9] Shichang Feng, Zhiquan Feng, and Liujuan Cao. 2019. Many-to-One Gesture-to-Command Flexible Mapping Approach for Smart Teaching Interface Interaction. *IEEE Access* 7 (2019), 179517–179531. https://doi.org/10.1109/ACCESS.2019.2957365

[10] Jean-François Jégo, Alexis Paljic, and Philippe Fuchs. 2013. User-defined gestural interaction: A study on gesture memorization. In *Proceedings of the IEEE Symposium on 3D User Interfaces* (Orlando, FL, USA) *(3DUI '13)*. IEEE, Piscataway, NJ, USA, 7–10. https://doi.org/10.1109/3DUI.2013.6550189

[11] Marit Korkman, Ursula Kirk, and Sally Kemp. 2007. *NEPSY-II: A Developmental Neuropsychological Assessment. APA PsycTests*. Psychological Corporation, San Antonio, TX, USA. https://doi.org/10.1037/t15125-000

[12] Gordon Kurtenbach and Bill Buxton. 1991. GEdit: A Test Bed for Editing by Contiguous Gestures. *SIGCHI Bulletin* 23, 2 (mar 1991), 22–26. https://doi.org/10.1145/122488.122490

[13] Joseph J. LaViola and Daniel F. Keefe. 2011. 3D Spatial Interaction: Applications for Art, Design, and Science. In *ACM SIGGRAPH 2011 Courses (SIGGRAPH '11)*. Association for Computing Machinery, New York, NY, USA, Article 1, 75 pages. https://doi.org/10.1145/2037636.2037637

[14] Frieder Loch. 2012. *Hierarchical Gestures: Gestural Shortcuts for Touchscreen Devices*. MSc. University of Twente, Twente, the Netherlands. https://essay.utwente.nl/61931

[15] Nathan Magrofuoco, Jorge Luis Pérez-Medina, Paolo Roselli, Jean Vanderdonckt, and Santiago Villarreal. 2019. Eliciting Contact-Based and Contactless Gestures With Radar-Based Sensors. *IEEE Access* 7 (2019), 176982–176997. https://doi.org/10.1109/ACCESS.2019.2951349

[16] 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

[17] Mykola Maslych, Eugene Matthew Taranta, Mostafa Aldilati, and Joseph J. LaViola. 2023. Effective 2D Stroke-based Gesture Augmentation for RNNs. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. ACM, New York, NY, USA, 282:1–282:13. https://doi.org/10.1145/3544548.3581358

[18] Zhenyao Mo and Ulrich Neumann. 2007. A Framework for Gesture Interface Design. *Journal of Multimedia* 2, 1 (2007), 1–9. https://doi.org/10.4304/jmm.2.1.1-9

[19] 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

[20] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding Users' Preferences for Surface Gestures. In *Proceedings of Graphics Interface* (Ottawa, Ontario, Canada) *(GI '10)*. Canadian Human-Computer Communications Society, CAN, 261–268. https://doi.org/10.5555/1839214.1839260

[21] Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of Pre-Designed and User-Defined Gesture Sets. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1099–1108. https://doi.org/10.1145/2470654.2466142

[22] 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

[23] Yosra Rekik, Laurent Grisoni, and Nicolas Roussel. 2013. Towards Many Gestures to One Command: A User Study for Tabletops. In *Proceedings of 14th IFIP TC 13 International Conference on Human-Computer Interaction. Lecture Notes in Computer Science, vol. 8118* (Cape Town, South Africa) *(INTERACT '13)*, P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, and M. Winckler (Eds.). Springer, Berlin, Heidelberg, 246–263. https://doi.org/10.1007/978-3-642-40480-1_16

[24] Quentin Roy, Sylvain Malacria, Yves Guiard, Eric Lecolinet, and James Eagan. 2013. Augmented Letters: Mnemonic Gesture-Based Shortcuts. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2325–2328. https://doi.org/10.1145/2470654.2481321

[25] 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)*. Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/1978942.1978971

[26] 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)*. Association for Computing Machinery, New York, NY, USA, 41–50. https://doi.org/10.1145/2396636.2396643

[27] 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

[28] 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

[29] Jean Vanderdonckt, Radu-Daniel Vatavu, and Arthur Sluÿters. 2024. Engineering Touchscreen Input for 3-Way Displays: Taxonomy, Datasets, and Classification. In *Companion Proceedings of the 16th ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (Cagliari, Italy) *(EICS Companion 2024)*, Michael Nebeling, Lucio Davide Spano, and José Creissac Campos (Eds.). Association for Computing Machinery, New York, NY, USA, 57–65. https://doi.org/10.1145/3660515.3661331

[30] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2012. Gestures as Point Clouds: A $P Recognizer for User Interface Prototypes. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction* (Santa Monica, California, USA) *(ICMI '12)*. Association for Computing Machinery, New York, NY, USA, 273–280. https://doi.org/10.1145/2388676.2388732

[31] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1325–1334. https://doi.org/10.1145/2702123.2702223

[32] 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

[33] 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)*. Association for Computing Machinery, New York, NY, USA, 131–138. https://doi.org/10.1145/2602299.2602316

[34] 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)*. Association for Computing Machinery, New York, NY, USA, Article 50,

3 pages. https://doi.org/10.1145/3531073.3534475

[35] 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. *Comput. Surveys* 56, 5 (2024), 128:1–128:55. https://doi.org/10.1145/3636458

[36] 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)*. Association for Computing Machinery, New York, NY, USA, 855–872. https://doi.org/10.1145/3357236.3395511

[37] 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

[38] 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* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[39] 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, Article 37 (2022), 54 pages. https://doi.org/10.1145/3503537

[40] Jialiang Yao, Terrence Fernando, and Hongxia Wang. 2012. A multi-touch natural user interface framework. In *Proceedings of International Conference on Systems and Informatics* (Yantai, China) *(ICSAI '20'12)*. IEEE, Piscataway, New Jersey, USA, 499–504. https://doi.org/10.1109/ICSAI.2012.6223046

[41] 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