

Iteratively Designing Gesture Vocabularies: A Survey and Analysis of Best Practices in the HCI Literature

HAIJUN XIA, University of California

MICHAEL GLUECK, Chatham Labs

MICHELLE ANNETT, MishMashMakers

MICHAEL WANG and DANIEL WIGDOR, University of Toronto

Gestural interaction has evolved from a set of novel interaction techniques developed in research labs, to a dominant interaction modality used by millions of users everyday. Despite its widespread adoption, the design of appropriate gesture vocabularies remains a challenging task for developers and designers. Existing research has largely used Expert-Led, User-Led, or Computationally-Based methodologies to design gesture vocabularies. These methodologies leverage the expertise, experience, and capabilities of experts, users, and systems to fulfill different requirements. In practice, however, none of these methodologies provide designers with a complete, multi-faceted perspective of the many factors that influence the design of gesture vocabularies, largely because a singular set of factors has yet to be established. Additionally, these methodologies do not identify or emphasize the subset of factors that are crucial to consider when designing for a given use case. Therefore, this work reports on the findings from an exhaustive literature review that identified 13 factors crucial to gesture vocabulary design and examines the evaluation methods and interaction techniques commonly associated with each factor. The identified factors also enable a holistic examination of existing gesture design methodologies from a factor-oriented viewpoint and highlighting the strengths and weaknesses of each methodology. This work closes with proposals of future research directions of developing an iterative user-centered and factor-centric gesture design approach as well as establishing an evolving ecosystem of factors that are crucial to gesture design.

CCS Concepts: • **Human-centered computing** → **Gestural input**;

Additional Key Words and Phrases: Gesture design, factors and methodologies, use-centered design

ACM Reference format:

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 (May 2022), 54 pages.

<https://doi.org/10.1145/3503537>

1 INTRODUCTION

Within **Human-Computer Interaction (HCI)**, *gestural input* most often refers to the user's finger, hand, arm, or other body part movements to express desired actions to computing systems.

Authors' addresses: H. Xia, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA; email: haijunxia@ucsd.edu; M. Glueck, Chatham Labs, 243 College St, Toronto, ON M5T 1R5, Canada; email: mike@chathamlabs.com; M. Annett, MishMashMakers, Newmarket, Ontario, Canada; email: michelle@mishmashmakers.com; M. Wang and D. Wigdor, University of Toronto, 40 St George St, Toronto, ON M5S 2E4, Canada; emails: michaelw.wang@mail.utoronto.ca, daniel@dgp.toronto.edu.



This work is licensed under a [Creative Commons Attribution International 4.0 License](https://creativecommons.org/licenses/by/4.0/).

© 2022 Copyright held by the owner/author(s).

1073-0516/2022/05-ART37

<https://doi.org/10.1145/3503537>

From early work exploring gestural input with a light pen [309] to research that proposed manipulating virtual objects using single finger gestures on a force and position sensitive screen [209] and more recent work that used mid-air gestures to enhance the creation of dynamic phenomena animations in virtual reality [25], the design, facilitation, and detection of gestural input has had a long history within HCI.

The methodologies used to design gesture vocabularies may be broadly categorized into three groups. With Expert-Led methodologies, seasoned gesture designers utilize their expertise, and past experiences with gestural interaction to create mappings between physical actions and system responses [55, 98, 120]. Expert designers can quickly create gesture vocabularies that are simple, socially acceptable, and ergonomic using the knowledge they have gained from watching users and interacting with gesture-based systems themselves. Although this is one of the most common methods used, the resulting gesture vocabularies have been criticized for lacking discoverability or intuitiveness [362]. User-Led methodologies, on the other hand, employ participatory elicitation techniques to generate gesture vocabularies based on user behavior or feedback about which gestures map most naturally to a given command [229, 258, 362]. The elicited gestures that result when one draws on the “wisdom of the crowd” may be intuitive and discoverable by a target user, but may lack other characteristics, such as not being differentiable during recognition, not fully leveraging the input capabilities of a given input device, or not being ergonomic during repeated, long-term use. The last group of methodologies, i.e., Computationally-Based methodologies, utilize algorithms to identify gestures that are easy for a system or interface to recognize and trivial for a designer to specify [16, 26, 106, 199]. While algorithms can generate a gesture vocabulary that is easy to recognize, it is difficult to algorithmically model other factors such as the learnability, transferability, or social acceptability of the generated gestures.

Although these methodologies all seek to leverage the unique expertise, backgrounds, and capabilities of different designers (i.e., experts, users, and computational models), the factors that a designer may wish to optimize often vary depending on the desired functionality, hardware capabilities, target users, or social context of an application. As a result, it is unlikely that the gestures designed using one methodology will fulfill all the requirements a designer may have. Furthermore, due to the lack of understanding, definition, and categorization of the relevant factors that are important to gesture design, many gesture vocabularies created using these methodologies contain gestures with socially unacceptable semiotics, gestures that are not safe for long-term use, or gestures that are not generalizable to other contexts or users [60, 220].

The goal of the present work is thus to provide an as-complete-as-possible listing of factors found to be important in the design of gestures, and to provide an index to the research that provides the field’s best-known methods for optimizing these factors. By identifying these factors, researchers and designers can develop a multi-faceted understanding of the many factors that are crucial to gesture design and understand the strengths and weaknesses of current gesture design methodologies along each of the identified factors.

To identify the factors crucial to the design of gesture vocabularies, over 1,600 gestural interaction research papers were surveyed and 288 papers within this body of literature that specifically focused on gesture design were synthesized further. This activity resulted in the identification, definition, and analysis of thirteen factors critical to gesture design including Situational factors (i.e., Context, Modality, Social Acceptability), Cognitive factors (i.e., Discoverability, Intuitiveness, Learnability, and Transferability), Physical factors (i.e., Efficiency, Complexity, Ergonomics, and Occlusion), and System factors (i.e., Feedback and Recognition). The holistic importance of the identified factors was explored by examining the three categories of gesture design methodologies from a factor-oriented viewpoint and highlighting the strengths and weaknesses of each methodology. Informed by these findings, we delineate two future research directions, including a potential

factor-centric gesture design approach wherein gesture designers identify, prioritize, design, evaluate, and refine the set of factors most related to their applications as well as establishing and maintaining an evolving ecosystem of gesture design factors. We believe these future directions can further the practices and the understanding of gesture design in both research community and industry.

2 DESIGNING GESTURES TODAY

Many methodologies have been proposed to assist in the creation of gesture vocabularies or gesture sets. Some methodologies rely exclusively on expert knowledge and experiences [6] or gather feedback and performance measures from users [6, 229] for gesture creation, whereas others emphasize how easy gestures will be to recognize [229] or use mathematical models and simulations [300]. Regardless of the methodology used, the general recommendation for the design of gesture vocabularies is that a resultant gesture vocabulary should contain interactions that are comfortable, efficient, easy to learn, and execute a user's desired actions [228].

2.1 Expert-Led Methodologies

The most traditional method that has been used to design a gesture vocabulary employs a top-down approach, wherein designers or researchers rely on their knowledge and experience with a target input device, end user population, or intended user actions to develop a gesture vocabulary. This process initially involves the identification and collection of a set of requirements for an application. Techniques such as developing scenarios [308] or use cases [131], performing a task analysis [81], wireframing [126], or storyboarding [313] allow one to understand end users' abilities, skills, environments, devices, and tasks. Using these requirements, a designer can then create a gesture vocabulary using gesture vocabularies that have been published by others [365], best practices and recommendations from the literature [376], commercial gesture vocabulary examples (e.g., Graffiti), metaphors or observations of analog interactions [119], or interaction designs that are based on a designer's experiences [354]. Once the vocabulary has been developed, a user study is often performed to evaluate the gestures using a given population, environment, device, or task. Factors such as learnability, discoverability, ergonomics, and social acceptability, among others, are typically of interest [194, 202, 289, 352, 364]. For example, Wickeroth et al. proposed a Gesture Usability Scale, which extended the System Usability Scale [53] to measure the usability of gestures based on their learnability, ergonomics, complexity, and recognition [352].

A variant of this traditional methodology proposed by Sturman and Zeltzer, utilized a five-stage method to assist designers in developing whole hand gestures [306]. With this method, a designer first determined if whole hand input was appropriate for their application by considering how natural, adaptable, and dexterous application movements needed to be. If deemed appropriate, the designer then consulted Sturman and Zeltzer's taxonomy of whole hand input to determine which types of gestures would be best suited to their application. Next, they decomposed the tasks to be performed into primitive actions and used these actions to decide which individual gestures would be best to use. These decisions could be made using the existing literature, the previous experience of a designer, or observations of how the hand is normally used in the intended, or similar, tasks. Next, the designer decided which device(s) would be used by the target population and evaluated the gestural vocabulary using these devices. After the evaluation, Sturman and Zeltzer noted that one may need to iterate through the five-stage process many times to ensure that the whole vocabulary met the identified requirements.

Barclay et al. systems-based view of gesture design proposed that designers work through a quantitative model that progressively measured the quality and functionality of individual gestures and the system as a whole [34]. First, the quality of a gesture was computed by combining

various weighted measurements of factors including fatigue, naturalness, duration, and accuracy. The results of each measure were then averaged to attain the overall quality of a given gesture. Then, it was recommended that a designer should derive a state transition model of each desired task within the target interface. Traversing through a state transition model, a designer should then identify the pathway that is most typical, using videotaped data of users working with the interface. This view then recommended that a designer compute a functional quality score using this pathway. Lastly, a designer should compute the weighted average of all tasks within the system as an indicator of the overall quality of the system. The resulting value could then be used to determine the impact of changing individual gestures within the system or task-gesture mappings.

Rossini proposed a parameter-based model inspired by communicative gestures [273]. After creating a gesture vocabulary (using an unspecified process), Rossini recommended focusing on the size of a gesture to determine its appropriateness for the intended population and context. Next, Rossini recommended deconstructing the gesture into its timing phases and attending to the position of the hands throughout execution. Once this was completed, Rossini recommended identifying the points of articulation of the gesture and the body space, or locus, involved in the gesture. Using such information, one should be able to determine the morphology of the gesture and use the morphology to evaluate and compare gestures using an end user population.

Although varying in the extent to which they abstract information and integrate the user within the feedback and evaluation phases, expert-led methodologies can enable designers to harness their knowledge and experience to quickly design and develop a gesture vocabulary. Dangers relating to unnatural hand or limb postures and repetitive movements can also be avoided as experienced researchers or designers should know to monitor for such characteristics when designing their vocabulary. As the designer is aware of the entire battery of tasks or functionality within the interface, they can also easily check for similarity issues within a gesture vocabulary to ensure that each gesture is distinct. A challenge with such methodologies, however, is that the evaluation methodologies that are used, or the extent to which users are part of the process, can hamper opportunities to understand the learnability or discoverability of gesture vocabularies. As the designer is often not the end user of the gesture vocabulary, it can also be difficult to determine how intuitive the derived gesture and functionality mappings are.

2.2 User-Led Methodologies

Methodologies that employ users as part of the design process have become very popular as of late. Wobbrock et al. presented a technique to create gesture languages that focused on how easy they would be to discover [362]. Inspired by participatory design, this elicitation methodology for **user-defined gestures (UDGs)** does not rely on the expertise of a designer. Instead, an experimenter showed everyday users the state of a user interface before and after potential input and then asked them to describe or perform the gesture that they would expect to yield the change in state. Rather than utilizing an interface, some experimenters described an outcome and asked users to mime or describe the gesture that would correspond to that outcome [362]. Once the experimenters collected responses from several users, they computed an agreement score to determine the “degree of consensus” between users with respect to the elicited gestures [200, 361, 362]. Gestures exhibiting higher agreeability were believed to be most appropriate for one’s application as they were identified by the most users as being an appropriate gesture for a given action.

The main advantage of such elicitation techniques is that the resulting gesture vocabularies are generated directly from user behavior and beliefs about the gestures that would most naturally map to a given action. By leveraging the “wisdom of the crowd”, it is argued that the elicited gestures would eventually converge toward the gesture that would most likely be initially guessed or tried by an everyday user. Furthermore, because participants devise and perform each gesture

without guidance, it is assumed that this methodology mimics users' first attempts at a gesture in the wild. The technique has also been reported to help achieve a "preferred" gesture vocabulary, since participants may consider their preferences, in addition to their first-blush reactions, before reporting their chosen gesture to experimenters [362].

This elicitation methodology has been widely adopted since the publication of Wobbrock et al. article at 2009, being used to create gestures for freehand TV control [212, 323, 383] and smart glasses [317], to unmanned aerial vehicles [249] and deformable displays [312]. Perhaps this widespread adoption is due to the academic and industry support for participatory and user-centered design, where the integration of users into a design process can help identify errors, ineffective solutions, suboptimal strategies, or unique requirements and challenges that could impact an interface before its final development or deployment [70]. We refer the audience to systematic review conducted by Villarreal Narvaez et al. for more details, which surveyed 216 gesture elicitation studies [334].

Preceding the work of Wobbrock et al., Nielsen et al. proposed the "human-based approach" to create gesture vocabularies [229]. With this approach, a researcher first identified the functions or features of an interface that gestures would be mapped to. The researcher then walked participants through a variety of scenarios and asked them to describe how they would invoke specific functionality or had participants use the interface within a Wizard-of-Oz paradigm. As the scenarios were videotaped, the footage could be reviewed to identify how consistently and frequently different gestures were used. To determine the gestures that constituted the vocabulary, a researcher attended to the frequency and duration of gesture performances, the internal forces a gesture posture created, as well as the effects that the movements could have on the hand. Once the gesture vocabulary was composed, this approach recommended evaluating the gestures for guessability, memorability, and ergonomics. Earlier iterations of Nielsen et al. human-based approach also recommended the use of guided drawing tests to match gestures to desired functionality and functionality to specific gestures [229].

Pyryeskin et al. utilized Wobbrock et al. approach within their work to design in-air tabletop gestures [258]. However, they noted that such a methodology does not enable one to evaluate users' performance of gestures or the ability for gestures to be recognized with limited hardware. They thus conducted a follow up study that used a target selection task with elicited gestures as input. Metrics measured during this experiment included gesture speed, gesture accuracy, and participant preferences. This two-study methodology enabled user input to be solicited and later refined per the requirements of the hardware that was available.

More recently, Wu et al. proposed a variant of Wobbrock et al. approach that integrated elicitation within a four-stage process [372]. In Stage 1, designers were encouraged to utilize semi-structured interviews to conduct a requirements analysis of an interface. Stage 2 utilized an elicitation study protocol that employed think-aloud methods, in addition to Wizard-of-Oz techniques (if appropriate), to determine two distinct sets of gestures for a given interface. Stage 3 used these gesture sets for a benchmark test that consisted of a degree of matching activity, memorability test, comfort test, and post-test questionnaire. The result of this third stage was a refined gesture vocabulary. Finally, Stage 4 made use of a personalized study design that employed the desired hardware and software to ensure that the gestures met the needs identified in the requirements analysis, the recognition algorithms were correct, and the gestures were evaluated for effectiveness, efficiency, and user satisfaction. The authors noted that at any point during the process, one may need to go back to a previous stage to make refinements.

Löcken et al. also proposed a four-stage process that integrated an elicitation methodology [191]. Unlike Wu et al. model, gesture designers first identified the usage context and functions of a system, but involved end users and other stakeholders in the process. Once the requirements were

identified, a participatory elicitation study was conducted to iterate on the identified functionality and collect gestures that could constitute the final gesture vocabulary. Designers then developed multiple gesture vocabularies from the collected gestures and used them in the last step to conduct a comparative evaluation between the proposed gesture vocabularies. This evaluation was thought to hone in on an optimal set of gestures, which may be a combination of gestures from different vocabularies. The authors recommended that this evaluation should include participants and elicit qualitative feedback to determine the optimal gesture vocabulary.

User-led methodologies enable researchers and designers to obtain first-hand feedback from their target population and understand potential discoverability and comprehension issues very quickly. As each of the elicitation variants including Nielsen et al. human-based approach emphasized the need for appropriate mappings between gestures and functionality, it is unsurprising that so many have flocked to this method [334]. The major difference between these variants, however, is that they differ in *when* users become engaged in the process (i.e., Löcken et al. invite them as part of the requirements analysis whereas Wu et al., Nielsen et al., and Wobbrock et al. begin engagement after the requirements analysis is completed). These variants also differ in the outcomes of the engagement (i.e., requirements, multiple gesture languages, empirical measurements, and so on) and in the overarching goals that are of interest (e.g., discoverability, memorability, comfort, and so on).

Although such elicitation methodologies have become popular, the creation of gestures by uninformed or novice users does not come without challenges. Placing the onus on end users to create gestures which may be used by a wide population nullifies the expertise of designers, who often take a holistic view of an user interface and consider many facets that are important to gestural input (e.g., the impact of repeated usage [229], interference between gestures within a vocabulary [381], the learnability of a vocabulary [186, 384], and the reliable detection and differentiation of gestures [381]). Users frequently design gestures that are appropriations of existing gestures or are limited by the technological capabilities of their past experiences, resulting in them neglecting other important factors such as the learnability, legacy bias, social acceptability, recognition, or the system feedback that would be provided [214].

2.3 Computationally-Based Methodologies

Rather than employing expert knowledge or users as part of the gesture design process, a growing number of methodologies have focused on developing algorithms or models that result in gestures that can be easily specified by a designer and are trivial for a system or interface to recognize.

2.3.1 Design-By-Demonstration. Early work on gesture interfaces focused on the design of tools to improve the specification, mapping, and recognition of stroke-based input using templates or demonstrations of gestures that were performed by gesture designers. Most often, pattern matching or learning algorithms were then employed to differentiate gestures based on their similarity to the provided templates or recognize gesture input.

For example, Rubine's GRANDMA toolkit [278] enabled a designer to record multiple instances of a target gesture and associate the gesture to a user interface element via gesture handlers. Once the gesture handlers were specified, the designer could evaluate the recognition accuracy of the gesture using various recognition algorithms. Similarly, with the **Gestural Interface Designer (GID)** [16], a designer could place user interface elements within an interface and map semantic actions (i.e., gestures) to each element. It is unclear how one would determine which gestures to use, but technical descriptions of how gestures would be recognized and represented within the system were provided. Similarly, systems such as Agate [172], the Gesture Design Tool [195], and iGesture [292] enabled designers to provide templates and examples of gestures to create

gesture vocabularies, as well as evaluate the performance of different classifiers via descriptive statistics, recognition results, and classification matrices. To facilitate this process, GHoST, proposed by Vatavu et al., utilized visualization techniques to aid in the analysis of gestures, by visualizing recognition errors and characteristics of gesture articulation patterns [327].

Several computer vision-based tools have enabled developers and designers to identify, track, and recognize hand, limb, or full-body gestures. Ashbrook and Starner's MAGIC system [26], for example, enabled designers to iteratively design motion gestures. Using MAGIC, a designer first recorded themselves executing a number of example gestures with multiple repetitions. They then browsed through the examples to view recognition results pertaining to distinguishability and consistency. A designer could then examine gesture recognition performance by performing the gestures that should or should not be recognized. In MAGIC 2.0 [160], Kohlsdorf et al. extended MAGIC to better account for the false positive motions that could occur in everyday life. By integrating indexable Symbolic Aggregate approximations, a designer could search through a database of everyday motions to determine if a given gesture could be accidentally triggered. Baytas et al. Hotspotizer [39] supported the declaration of locations where gesture input and recognition should be more precise and areas where it should be less well-defined to allow for more fine and coarse-grained recognition and tracking. Other tools such as Paper-Mache [158], crayons [88], and Eyepatch [207] also enabled a user to demonstrate a motion and test different recognition classifiers iteratively, albeit within a specialized interface that did not enable new recognition algorithms to be authored easily. The DejaVu system [148], however, was built within an IDE to enable gesture developers to visually and continuously monitor real-time gesture input, edit recognition algorithms, and inspect recognition results.

Although pattern matching with templates or using examples of gestures can be viable methods to quickly evaluate recognition algorithms, projects such as Gesture Coder [198] and Gesture Studio [199] supported the declaration of gestural input via high-level language-based constructs. Gesture Coder enabled multi-touch gestures to be programmed by demonstration within the Eclipse IDE. Its successor, Gesture Studio, further enabled a designer to record a demonstration of a desired gesture, revise and edit the demonstration, create high-level behaviors that were composed of multiple gestures or movement clips, and attach actions to the behaviors and gestures.

The aforementioned tools support designers in quickly capturing gestural data and prototyping recognition algorithms, however, these tools only evaluate recognition performance. Quill [196], however, enabled designers and developers to view accuracy information as well as empirical data and visualizations about how similar a gesture was to existing gestures in a vocabulary. This software also provided active feedback, i.e., hints and advice, that a designer could use to eliminate similarities between gestures. Such feedback could be an important indicator when determining how difficult gestures would be for a user to recall.

Other projects have focused on easing the development of gestures for tangible, mobile devices. The aCAPpella system [106], for example, enabled everyday users to demonstrate context-aware behaviors for recognition, extracted and collected relevant sensor information, detected events that occurred within the sensor stream, and invited the user to annotate events corresponding to the behaviors they wished to support. The system then utilized this information to train an interaction model and incorporated it within a context-aware behavioral recognizer. With Exemplar [111], a designer could connect sensors to a PC and record examples of themselves performing desired motions. The streaming data was then visualized using small multiple views, which enabled designers to interactively evaluate different filters using the live data stream. Once satisfied, the designer could indicate which elements of a motion signal they wanted to recognize and the system would compute how accurately the signal would be recognized amongst the other recorded samples. Mogeste [245] also enabled designers to create new gesture vocabularies by recording mobile phone

movement and motions, selecting elements of the movement that defined a target gesture, and testing the recognition using built-in classifiers. In a slightly different vein, a suite of applications developed by Kim and Nam [152] and Kim et al. [153], enabled designers and everyday users to prototype interactive applications that used touchscreens, cameras, or sensors-embedded objects. Designers could demonstrate a target gesture and have various motion signals captured by embedded sensors. The designer could then specify the state-transition diagram of the resulting sensor streams using an interactive visual markup language to define their desired gestures.

2.3.2 Model-Driven Development. Aside from design-by-demonstration approaches to gestural design, other notable projects have employed the use of declarative and theoretical models of human behavior [78, 180, 184]. Leiva et al., for example, postulated that gesture vocabularies could be made by generating a large amount of gestural variants that were derivatives of a single example gesture [180]. Using the kinematic theory of rapid human movement, they argued that such variants could be used to improve recognition and negate the need to recruit participants to evaluate gesture languages. With their system, Gestures Go Go, designers could iteratively refine synthesized gestures by removing variants they found unsuitable and then generating more variants. The designers could also select and export a gesture recognizer (from a set of built-in recognizers based on the programming language of their choice) and a synthesized gesture set, to reduce the development time and efforts of gesture-based applications. The generated gesture vocabulary could then be evaluated using various methods, such as Leiva et al. Omnis [185], the GATO [183], or the KeyTime technique [181].

Work by Stern et al. viewed gesture vocabulary design as a multi-objective optimization problem [300]. With their approach, a designer identified the tasks, commands, and gestures that should be used in a system. They then conducted user experiments to determine how intuitive, stressing, and how frequently used the gestures would be. These results were then used to compute an intuitive matrix, comfort and stress matrices, and a reduced gesture language matrix. The gesture language matrix was then optimized to minimize the size of the gesture vocabulary and maximize recognition accuracy. The resulting optimization was combined with the intuitive, comfort, and stress matrices to match gestures to commands such that Pareto optimal solutions were obtained. From the generated solutions, the designer could then select the gesture vocabulary that best met their preferences.

Hochreuter's LemonGrasp [123] enabled designers to specify and design multi-touch gestures using tunable Manipulation-Attributes, such as speed, number of fingers, and inertia. This interface enabled designers to define gestures and the state changes (i.e., feedback) that should accompany the performance of each gesture. With Gestit [296], Spano et al. proposed a method for creating gesture interfaces that utilized compositional and declarative gesture definitions to define interaction. The framework enabled developers to use declarative notations to define their own gestures or combine existing gesture elements into gesture sequences. The notation was then used to associate gestures to specific UI views and elements and to recognize the gestures. The SNAP programming tool [155] also followed a programming-by-demonstration notion, but used a conceptual model that viewed gestures and motions as "triggers", devices that provide feedback as "objects", and device feedback as "responses". The software then enabled designers to specify interactions by composing these elements.

Each of these systems ensures that the resulting gestures are easy to recognize and create, rather than focusing on user-facing aspects of gesture design. Many of these systems are also targeted toward everyday users, but it is this population that may be unaware of the challenges that repeated execution, social acceptability, or environmental context may have on the design and use of gesture vocabularies. This neglect can be detrimental to users if, for example, the resulting gestures

are difficult to remember, cumbersome or awkward to execute, or illogically mapped to functionality. Such systems would be better suited as components within a larger gesture design process that incorporates processes to ensure gestures focus on these important factors as well.

2.4 Summary

Regardless of the methodology chosen, the process of creating gestural vocabularies remains a cumbersome, time consuming, and error-prone endeavor. While each aforementioned methodology has use cases or situational requirements that it is best suited for, what one method excels at, another fails or does not even consider (Table 1).

In the case of user-led methodologies, despite using gestures everyday, many users are often unaware of the facets of gesture design. On the other hand, when experts design gestures, their expertise and concern for environmental or technology factors such as recognition rates or desired modalities can hamper their abilities to design gestures that are discoverable or learnable. Gesture vocabularies that are designed via modeling or by demonstration are only as good as the underlying models used to generate, model, or evaluate them. Systems such as Gesture Go Go, LemonGrasp, Gestit, and SNAP offer though-invoking visions of future gesture design processes because they suggest that the best gestures may be those that humans cannot even conceive of.

More recently, some researchers have proposed combining or chaining different approaches for gesture vocabulary design. For example, building upon of the concept of constructing complex gestures from primitive gesture components [190], Delamare et al. explored user-defined combinations of gesture primitives which were previously designed by experts in the context of smart TV interaction. Vuletic et al. explored the chaining of User-led and Expert-led approaches to design gestures for conceptual design applications [341]. Within their approach, researchers would perform a gesture elicitation study (i.e., User-led methodology), and then professionals of varied backgrounds would be asked to evaluate the resulting gesture set (i.e., Expert-led methodology) so that the gesture vocabulary could be expanded or pruned as necessary to be suited to a larger audience and be easier to learn. Vuletic et al. found, however, that after going through this exercise and implementing the gesture set, it would necessary for researchers to undergo additional rounds of evaluation of the resulting gesture set.

While an abundance of methodologies and systems have eased some of the processes required to create gesture vocabularies, they also call for a cohesive, clear set of criteria or factors that can be used to design and evaluate gestures. Identifying the factors that constitute an “optimal” gesture vocabulary is essential because they could further enable a dissection of the strengths and weaknesses of existing and future methodologies and systems. In addition, they would enable gesture designers to design, evaluate, and iterate on gesture vocabularies using factors that are most relevant to their use cases, rather than being constrained to the subset of factors that a methodology is currently capable of evaluating.

3 FACTORS IMPLICATED IN GESTURE DESIGN

To understand the multitude of factors implicated in the design of gestures, a survey of the literature on gesture design was conducted. As the terminology used to describe the facets of gestural interaction has evolved over many decades, it has resulted in a fragmented lexicon. This section synthesizes the literature on gesture design and provides a holistic and comprehensive understanding that designers, developers, researchers, and the field have been missing.

3.1 Methodology

To collect a corpus of representative literature, specific terms including “gesture design”, “design of gestures”, “gesture tools”, “gesture set design”, “gesture toolkit”, “gesture factors”, “gesture

Table 1. Summary of the Differences between the Gesture Design Methodologies Proposed and used within the Literature: Expert-Led, User-Led, and Computationally-Based

Characteristic	Expert-Led	User-Led	Computationally-Based
Empirical evaluation of vocabulary possible?	●	○	●
Users involved in the design process?	○	●	●
Users involved in the validation process?	●	●	●
Potential for vocabulary to be biased toward past experiences?	●	●	●
Long-term validation process?	○	●	○
Single stage or multi-stage process of design?	Both	Both	Single

Legend: ● – Yes; ● – Maybe; ○ – No.

software”, and “gesture vocabulary design” were queried on the ACM Digital Library, IEEE Xplore Digital Library, Google Scholar, and Microsoft Academic Search. This resulted in 3,352 publications being collected as of April 2021. After removing duplicate entries (i.e., 1,277 publications) and publications that were not in English (i.e., 144 publications), 1,931 publications remained. To identify publications that focused on gesture design, we define the inclusion criteria as that the publication must focus on designing gesture vocabulary for human–computer interaction scenarios, devising methodologies for gesture design, or studying the aspects that can affect gesture performance. In total, 1,643 publications were excluded as they focused on unrelated themes such as American Sign Language learning, computer music, gesticulation within the fine arts, gesture recognition algorithms, and so on, or because they republished the same results in multiple venues or utilized a prior publication’s gesture vocabulary within a new domain instead of iterating on or developing a new gesture vocabulary. Finally, a collection of 288 publications remained, which covered the design of both 2D gestures using the fingers or hands, as well as 3D gestures using the arms, head, or feet.

Using the search terms as a starting point, the best practices, recommendations, experimental results, and future work themes discussed in each publication were classified using an open coding method [305]. Themes were (re)classified or aggregated as new factors emerged (e.g., *learnability* was found to be referred to as *Fast Learning*, *Memorability*, *Recall*, *Ability to Recognize*, *Similarity*, *Fits Well with its Associated Function*, *Systematic Understanding*, *Appropriate Mappings*, *User Uptake*, *Action Matches Function*, and so on, so all of these “factors” were aggregated under the umbrella term *learnability*). This process underwent many iterations and resulted in the identification of thirteen factors crucial to the design of gesture vocabularies (Figure 1). For readability and improved organization, the factors are grouped into four categories, based on the interconnected relationships between factors, e.g., those related to the interactive system itself, such as *feedback* and *recognition* were grouped into *System Factors*. Herein, each factor is defined and described using relevant examples and findings from the literature.

3.2 Situational Factors

Governing the development of any gesture vocabulary are the overarching goals that a designer is trying to achieve and constraints they must contend with. Within the context of gesture design, three themes important to situational constraints and opportunities emerged, including the *context* where the gestures will be performed, the *modality* used to perform the gestures, and the *social* factors that may influence a user’s performance and willingness to perform the gestures. Most

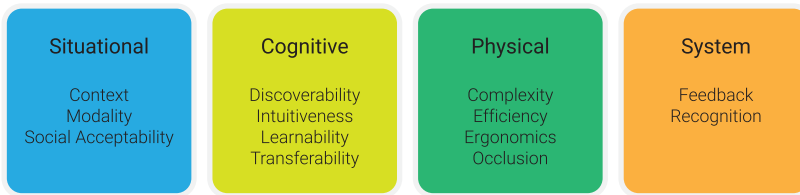


Fig. 1. The 13 factors that were found to be essential for gesture design are grouped according to those that were situational, cognitive, physical, and system-based.

often, these factors are found within research that employed Expert-led methodologies, as it is experts and designers who decide the domain and scope of the activities associated to the gestures.

3.2.1 Context (i.e., Users, Environments, and Tasks). As defined by Dey, *context* describes any information that characterizes the situation where a gesture will be performed (i.e., people, places, and devices) [79]. The literature review demonstrated that the term *context* is used infrequently, whereas terms such as users, environment, or task¹ were more common. If gestures are performed in a vehicle while driving, for example, a designer should identify the limitations of sitting in the driver's seat and the risks associated with divided attention, as well as seek opportunities that minimize the movements of hands away from the wheel or the eyes away from the road [15, 86, 250, 252]. If children [61, 69, 189, 328], seniors [61, 303, 304], or those with motor [46] or visual impairments [142, 144] will be performing the gestures, it is important to identify the unique challenges that such a population may incur. For example, if a designer is creating a gesture vocabulary for elderly users, they should consider how the physical limitations and memory issues that often occur with age may influence the recall and performance of the gestures [303, 304].

Before one can evaluate the degree to which their gesture vocabulary meets or exceeds each of the challenges or opportunities that the context imposes, designers must determine the requirements of their gesture vocabulary, which is often done before the gesture vocabulary is created. However, as the survey demonstrated, none of the gesture vocabularies created to date weighed people, places, and devices equally. Thus, a requirements evaluation can serve as a sanity check for designers, ensuring that they are aware of all the challenges that their gesture vocabulary may encounter.

To determine the requirements of their gesture vocabulary, one could use techniques such as working through scenarios [308], storyboarding [313], use case development [131], task analysis [81], or wireframing [126]. Robertson and Robertson also proposed the **Volere Requirements Specification Method (VRSM)**, which is used within engineering, organizational management, and product design to identify the importance of preconditions of a product [272]. It contains 27 types of requirements that include, among others, legal, environmental, cost, scope, and performance. To identify a requirement, one would complete a template scorecard (i.e., Requirement Shell) to articulate the rationale and importance of the requirement, as well as identify the consequences that could occur if the requirement is not considered. This process lends itself to the holistic and extensive understanding of gesture requirements.

¹Reviewed literature included [3, 6, 10, 15, 23, 28, 31, 33, 36–38, 45, 47, 48, 52, 54, 55, 60–62, 64, 65, 68, 69, 73, 86, 90, 91, 93, 96–98, 100, 102–104, 109, 110, 114–119, 122, 124, 132, 134, 138, 142, 142, 144, 151, 157, 160, 161, 167, 169, 170, 175–177, 179, 186, 187, 189, 191–193, 201, 206, 208, 213, 215, 218, 219, 227, 234–238, 240, 242, 244, 246, 250, 252, 253, 255, 258, 261, 264, 267, 271, 280, 281, 284, 288, 290, 291, 294, 295, 301, 302, 304, 311, 314–316, 321, 323, 324, 329, 331, 333, 338, 342, 344, 345, 347, 349–351, 354–356, 362, 364, 366–368, 371, 372, 375, 382, 384, 386].

3.2.2 *Modality*. According to Nigay, interaction *modality* with HCI has been discussed at multiple levels of abstraction. For example, a modality could be specified in general terms as “using gestures” when comparing with modalities such as speech and keyboard, or more specifically as “using finger gestures on a multi-touch screen” when comparing with arm or body gestures [72]. Today, a common level of abstraction has been the physical devices (e.g., touchscreen devices) and the interaction languages (e.g., a set of 2D gestures) that a user would employ to achieve their goals [232]. Modalities have also been referred to as *enabling technologies* [146] or *tools* [113] within the literature.

Within the context of gesture design and for the purpose of distinguishing the many different types of gestural interaction, we employ the lowest level of abstraction for modality, which refers to the sensory input channel used to perform a gesture and the output channel that a system uses to provide feedback to the user [72]. Designing for specific modalities has been, unsurprisingly, a very popular topic to explore.²

In terms of input, modalities come in a variety of forms: fingers [265, 302, 303, 315, 374], hands [187, 241, 244, 267], feet [134, 140, 156], nose [247, 385], the entire body [69], implements such as styli [63, 316], and tangible objects such as steering wheels [15], controllers [64], or remote controls [372]. Each input modality has a series of contexts that it works best for. Touch, for example, is best for the direct physical manipulation of objects and is ubiquitous in mobile settings, mice are preferred for their precision and accuracy, styli best mimic the nuances of inking and sketching, and the keyboard is one of the most efficient input devices. The reader is referred to Karam’s review of gestures in HCI for more specific details [146]. Such diverse input options, however, introduce constraints for designers, not only in terms of which modality to support, but also if and how execution will differ between modalities (e.g., finger and pen [316], or finger and arm [89]). Although finger and arm-based gestures are the most popular, gesture designers also need to be aware that alternative, less common input channels such as the feet, legs, elbows, and so on may be better suited for input when, for example, the hands are occupied or attention is divided (e.g., [112]). Note that work by Köpsel et al. indicated that the input modality chosen may play less of a role in gesture performance than its context [162].

In terms of output, the technology that is available in an environment can influence how gestures are executed and users’ responses to them. Many investigations have been undertaken to understand which gestures are best matched to a given technology (e.g., tablets [259], televisions [323], mobile phones [30, 48, 363, 390], multi-touch displays [173], and so on). For example, if only a large screen is available, users will have a tendency to make larger gestures, which will take longer to perform [127]. If content is being transferred between displays of different sizes, users have been found to first gesture to make the content smaller and then gesture to transfer the content [170]. Devices that have a smaller form factor such as tablets or mobile phones enable users to gesture *with* and *on* them, both when stationary and on the go [48, 316]. Larger devices such as computer monitors and wall-sized displays do not offer such opportunities. Differences have also been found in the smoothness and number of gesture repetitions users will perform while using tangibles and physical objects versus virtual proxies and digital content [217]. Even the techniques used to detect gestural input, e.g., optical versus surface acoustic wave, have been shown to impact the nature of gestural movements [97]. In other cases, environmental factors may limit the technology that can be used. For example, in-air gestures are often preferred over touch-based gestures for hygiene reasons in the kitchen [98] or during surgical procedures [274].

²See references including [6, 15, 17, 28, 31, 37, 40, 47, 48, 54, 55, 68, 69, 73, 85, 86, 90, 93–97, 100, 108, 119, 122, 125, 132, 134, 138, 142, 147, 157, 161, 166, 176, 186, 187, 191, 201, 204, 208, 213, 218, 227, 235, 236, 240, 241, 244, 250, 257, 265, 266, 280, 282–284, 288, 300–304, 310, 311, 315, 323, 324, 331, 335, 338, 343–348, 350, 367, 371, 376, 377, 382, 385].

The utility and usability of modality-specific gesture vocabularies are often evaluated via Expert-led methodologies that involve end users or close approximations of them. Experts typically work with a specific modality, or a number of different modalities to compare performance and often have user study participants' movements, behavior, and opinions of the modalities observed and measured [42, 66, 92, 377]. If gestures performed using a specific modality are found to require more time and effort to learn, a designer may need to modify or redesign the gestures. A designer also may need to switch to an alternative modality, but this could also result in the redesign of the entire gesture vocabulary. We recommend that designers select a modality taking a holistic consideration of the identified requirements within their desired context.

3.2.3 Social Acceptability. *Social acceptability*, as articulated by Williamson, refers to the social and cultural factors that affect the user experience when a gesture is performed [358]. Such factors can include the location where the gesture will be performed, the presence and composition of the audience in this location, the age of the user, the body parts and areas that the gesture requires, the size and duration of the gesture, and so on.³ Within the literature, social acceptability has also been referred to as *social comfort*, *cultural acceptability*, and *cultural transparency*. Although interaction is largely considered to be an individual experience, acceptability includes both how the user feels while performing the gestures (i.e., the user's social acceptability) and how nearby spectators will perceive the gestures (i.e., the spectator's social acceptability [269]). Thus, the perceptions and reactions of others to the performance of gestures are important aspects of gesture design that cannot be overlooked.

Koelle et al. identified four design strategies employed in the HCI literature to improve the social acceptability of interaction techniques, which include using subtle and unobtrusive interaction to avoid negative attention, avoiding suggestiveness and misinterpretation, designing interactive devices that are accessory like and of familiar styles, as well as making interaction more candid, transparent, and observable [159]. These design strategies, however, do not always conform with each other. Pohl et al. also found that hiding and deception have been employed to enable subtle and unobtrusive interaction in social settings [256]. For example, Anderson et al. proposed using deceptive and illusory techniques informed by principles of magic to improve subtlety of interaction with digital devices to avoid negative effects on in-person interaction.

Rico and Brewster argued that acceptability is not purely a binary measure, but rather occurs along a continuum [269]. Some gestures may be more appropriate to perform in certain locations (e.g., in an elevator versus on a street corner) or with certain groups of spectators (e.g., with family versus strangers). The social acceptability of a gesture can be further influenced by one's need to repeatedly perform movements due to poor recognition, if the gesture needs to be large and cannot be performed subtly, or if the gesture employs mimetic, instead of alphabetic, movements [87]. Cultural norms also play a role, as hand and bodily gestures from one culture may be offensive or inappropriate in another. When users deem gestures to be less acceptable or appropriate to perform, they will be more hesitant and less likely to utilize them, or if they do perform them, will feel uncomfortable [2, 82].

Best practices to achieve culturally- and socially-resilient gesture vocabularies recommend asking end users to perform gestures in the wild or in a laboratory. Participants are then asked via think-aloud protocols or questionnaires about how they feel while or after performing the gestures [2, 5, 82, 99]. As such methodologies require users to perform gestures in public settings or around people they do not know, "if users decide to interact by gestures (in the wild), they

³See references including [2, 5, 14, 82, 87, 99, 134, 159, 211, 240, 256, 263, 269, 288, 299, 335, 350, 358].

implicitly categorize this interface as socially acceptable” [99]. This enables a designer to evaluate social and cultural acceptability alongside other factors without incurring additional costs.

Gestures have also been recommended to be small, unobtrusive, unnoticeable [211], physically comfortable, enjoyable, similar to everyday movements or gestures used with existing technologies, and not interfere with communication [358]. As suggested by Rehfeld et al., assigning gestures to functional mappings such that gestures are universal across cultural contexts can make it easier for users to learn and remember gestures [263].

Lastly, repeated exposure to, and performance of, some gestures can result in them changing the cultural zeitgeist and becoming more socially acceptable over time [299]. Examples of now-culturally appropriate gestures include the swipe gesture that users performed on the temples of Google Glass, waving toward a Kinect sensor, and using foot gestures to interact with virtual spaces [134]. It is important for designers to remember that as certain gestures become more or less appropriate over time, it may be necessary to iterate on their gesture vocabularies later.

3.3 Cognitive Factors

In addition to situational factors, another theme that emerged was the importance of cognitive factors to gesture design. A number of studies within the literature have focused on the importance of the *discoverability* of the gesture vocabulary, how *intuitive* it is to use, how *learnable* the gesture vocabulary is, and how *transferable* the gesture vocabulary is to other domains or contexts. Each of these factors is affected by the user’s cognitive understanding of the system and the motions they should be performing. Cognitive factors are often found to be important with User-led, and occasionally Expert-led, methodologies [17, 121, 362, 364, 378].

3.3.1 Discoverability. The quality of gestures that enables a user to access intended referents of those gestures despite a lack of knowledge about the gesture is often referred to as the *discoverability* of a gesture [361]. This factor has also been termed *guessability*, *approachability*, *self-revealing*, and *metaphorically or iconically logical towards functionality*⁴ and is most often considered during User-led methodologies. Regardless of the terminology used, in some contexts it may be crucial that a gesture or gesture vocabulary is discoverable. For example, if one is required to perform a gesture to initiate interaction with a kiosk that they are encountering for the first time, it is important to have a discoverable gesture vocabulary so they can quickly access system functionality, rather than become frustrated and walk away.

As proposed by Wobbrock et al. [361, 362], to determine how discoverable a gesture is, a gesture vocabulary designer can compute the agreement rate or agreeability of a set of gestures using various methods [9, 326, 331]. The agreement rate determines how similar the gestures elicited from different participants are to each other, and thus how easy the gestures would be to guess and perform in the wild [334, 340, 361]. As an alternative to this, Seto proposed the use of an in-the-wild observational paradigm, which used video data to measure how frequently gestures were performed by uninformed users in the wild and suggested that frequency could be used as a measure of discoverability [289].

If a gesture set or individual gesture is deemed undiscoverable, many techniques can be used to improve its discoverability. First, the output presented to users could be changed to provide pre-execution information about the gestures a system supports (i.e., self-revealing gestures [38] or visual and audio cues ([20, 381])). For example, using the hover state, a gesture cheat sheet, or an animation or video of the gestures that could be performed could be displayed [18, 357]. If visual feedforward and feedback cannot improve discoverability, a designer could also make use of an

⁴Reviewed references included [1, 9, 31, 58, 110, 175, 271, 289, 326, 331, 332, 334, 340, 345, 360–362, 378, 385, 386].

action-outcome mapping that mimics common actions (e.g., like Hinckley et al. cutting and pinching gestures ([121]). Piggybacking on an already familiar or discoverable gesture is another way of increasing the discoverability of functionality [378]. To reduce the need for users to discover and learn a new set of gestures when interacting with a new application or interactive device, Vatavu proposed nomadic gestures, which allow users to reuse their preferred and practiced gestures to interact with newly encountered devices [322].

As more everyday users become collectively required to use and acquire gestures to complete day-to-day tasks, the higher the likelihood that gestures that were not previously discoverable become discoverable. Therefore, in some cases, it may be appropriate for developers to utilize widely adopted gesture vocabularies for common interaction tasks.

3.3.2 Intuitiveness. In 1994, Raskin stated that *intuitiveness* refers to the degree to which a gesture makes use of transferable and existing knowledge [260]. Years later, Jacob et al. furthered this by recommended that newly proposed interaction techniques should draw strength from users' pre-existing knowledge of the everyday, non-digital world [130]. For example, swiping to the left to flip a virtual page could be considered an intuitive gesture because this action mimics the physical movements made when turning a page [343]. Intuitiveness has also been referred to as *appropriateness, familiarity, and necessity*.⁵ Intuitiveness is an important factor to consider because it ensures that the “*gesture afforded by the interface design is aligned to how users expect to provide input*” [173] and that there is an appropriate “*cognitive association between [the] command or intent, and its physical gestural expression*” [300]. If the mapping between the action or movement does not match the resulting functionality, or it is hard for the user to draw parallels or build upon the existing knowledge they have, they will perform the incorrect gesture or not perform any gesture at all. Note that a gesture may be intuitive without being discoverable (e.g., prior to 2007, novice touchscreen participants would rarely guess the “pinch to zoom” gesture, though once shown by an experimenter, would recall it flawlessly [357]).

Many methods can be used to measure how intuitive a gesture is. Nielsen et al., for example, proposed two approaches [229], (i) a bottom-up approach similar to Wobbrock et al. UDG methodology [362] that presented actions or functionality to users and asked them to create a suitable gesture and (ii) a top-down approach that showed users a gesture and asked them to select the functionality that should be associated to the gesture from a candidate list. With these approaches, it is common to ask novices [58, 101, 102, 115, 128, 300], experts [38], or members of a design or research team [100] to comment on intuitiveness via questionnaires, interviews, video transcriptions of the think-aloud words used by users while performing the gestures, and so on. As proposed by Agarwal and Prabaker, one could also empirically measure intuitiveness by the computing the differences in task duration between experts and novices [3].

If gestures are not found to be intuitive, it would be beneficial to change the mapping between the gesture and the outcome, making use of a mapping that assigns the gesture to more cognitively similar actions or mimic natural movements [121, 381]. It is also important to note that the repeated exposure and performance of gestures in one context may make them feel more intuitive in another context. For example, Wobbrock et al. found that the traditional desktop user interface metaphor was deeply rooted in the participants' mental model and significantly influenced gestures they created for multitouch surfaces.

3.3.3 Learnability. The notion of *learnability* refers to the ease with which new users can begin effective interaction and achieve maximal performance with a gesture [83]. Learnability is one of

⁵See references including [1, 3, 4, 38, 54, 58, 73, 86, 98, 100–102, 115, 116, 128, 130, 134, 157, 169, 170, 173, 176, 177, 179, 191, 206, 224, 229, 235–237, 244, 250, 252, 255, 260, 282, 294, 295, 300, 301, 319, 342, 350, 371, 384].

the most popular factors within the literature and is often referred to as *fast learning, memorability, recall, recognize, similarity, fits well with its associated function, systematic, chunking, mapping, uptake, or matching*.⁶

Learnability and intuitiveness are complementary, as the more intuitive a gesture is, the easier it is to learn. However, intuitiveness refers to the mapping between a movement and outcome, whereas learnability focuses on how this mapping, in addition to other factors, enables one to recall a gesture. If a user is never able to learn a gesture vocabulary, it will not be used. Learnable gesture vocabularies are beneficial because they negate the need for extensive training.

There are many methods that a designer can use to determine how learnable a gesture or gesture vocabulary is. Computing the number of gestures a user can remember is one method to determine learnability [21, 34, 50]. Wobbrock et al. proposed using techniques such as counting the gestures required for a unit action (i.e., gestures per character for text entry) or measuring the time it takes to perform each gesture [364]. These measures assume that less learnable gestures will have higher numbers of gestures per character and longer task times.

If training will be provided to a user before they will perform a gesture, then learnability may not be as important. As there have been many investigations into the best way to teach gestures, the reader is referred to Anderson and Bischof [13], Bau and Mackay [36], Bragdon et al. [50], or Yee [381] for more details.

3.3.4 Transferability. Perkins and Salomon defined *transferability* as the degree to which the learning of a gesture can enhance or undermine its performance in another context [248]. Within the literature, this has also been termed *generalizability, transfer to expertise, or adaptability*, and has been applied to transferring learned gestures to other hands [17], modalities [294, 295], applications or devices [368], user populations [182, 325], and to expert performance [13, 34, 50, 93, 139, 384].⁷ For example, [325] framed interactive gestures as a type of knowledge that can be characterized by the information of the gesture movement, having an understanding of the situation where the gesture is performed, and the experience and skills of the user that leads to efficient execution of the gesture. [325] further proposed that gesture knowledge should be easily transferred to situations requiring different gesture articulation, sensing environments, and interpretations of users' preferences.

Transferability is similar to learnability [67] in that to use a gesture in a different context, a user must first learn a gesture and then later recall it on demand. It is different, however, in that certain gestures lend themselves better to transferring to a different context or situation. Consider transferability to expert performance as an example. Marking Menus [171] gestures enable users to gradually improve the efficiency of gesture performance and robustness to error through repetitive executions. If gestures can be modified to chain off each other or utilize chunking, transferability can be improved [56]. This may, however, result in a gesture being composed of entirely new movements. Training and tutorials for gestures can also facilitate progress toward expert performance.

Transferability is often only considered when Expert-led methodologies are used. To measure the transferability of a gesture or entire gesture vocabulary, one can use a retention and transfer paradigm from the motor learning literature [13, 17]. As suggested by Silpasuwanchai and Ren [294, 295], a designer could also ask users from the target population to comment on other body parts that could be used to generate a given gesture. Given that this factor is dependent on the user's cognitive ability to remember and execute motor movements, it is imperative that users

⁶As suggested by the following: [1, 6, 9, 10, 13, 19, 21, 24, 33, 34, 36–38, 49, 50, 55, 83, 93, 94, 98, 101, 104, 116, 137, 169, 175, 178, 186, 191, 192, 197, 208, 224, 244, 250, 266, 277, 283, 284, 288, 294, 295, 300, 318, 324, 326, 330, 333, 342, 344, 347, 356, 364, 366, 368, 371, 372, 387].

⁷See references including [13, 17, 34, 50, 56, 67, 93, 139, 171, 182, 248, 294, 295, 325, 368, 369, 384].

are involved in the design or evaluation of a gesture vocabulary that is intended to be transferable [13, 17].

3.4 Physical Factors

The third theme that emerged within the literature were those factors that were influenced by the physical movements that users make. These factors largely relate to the *efficiency* with which gestures can be performed, the *complexity* of the movements, how ergonomically *appropriate* the movements are, how *natural* the movements feel, and the degree to which the interface or feedback provided to the user is *occluded* by the motions of the gestures.

3.4.1 Complexity. The Oxford Dictionary defines *complexity* as the state or quality of being intricate or complicated [243]. When referring to gestures, this implies there is a certain degree of difficulty to the movements that must be remembered or performed. Complexity may also be referred to as *mental load*, *performability*, *simplicity*, or *attention*.⁸ The gestures that a user is required to perform should have the appropriate level of complexity given the task at hand. When using a mobile phone, for example, the gesture performed to unlock the phone, which may be done hundreds of times a day, should be as uncomplicated and simple as possible so that the user can perform it in a variety of situations with relative ease. In other situations, however, it may be appropriate to require that a more complex gesture be performed. If one is using an in-air gestural interface to direct the movements of an endoscopic surgical robot, for example, the complexity of the movements should match those the surgeon would normally perform during non-robotic surgery to ensure that the robot performs the correct maneuvers and that gestures are not accidentally identified.

Measuring the complexity of a gesture requires one to critically analyze the constituents of the gesture movement, such as the number of steps, repetitive patterns, or distinct, discrete movements each gesture requires [285]. Production time measurements and efficiency estimations with the models used to evaluate efficiency can also indicate the complexity of a gesture. As recommended by Yee, one should also manually check if a gesture requires unusual movements that increase the effort one would need to exert [381]. As a majority of the methods to determine complexity require empirical measurements, it is quite common for research concerned about complexity to utilize Expert-led or Computationally-based methodologies to create gesture vocabularies. For example, Leiva et al. proposed Omnis Prædictio, a generic technique that can provide user-independent estimations of the many numerical features of gestures, such as production time, speed, and curviness [185].

Complexity often influences other factors as well, e.g., a complex gesture may be hard to learn because it is difficult to use and not discoverable [179]. As advantageous as it is to have a simple gesture vocabulary, in some contexts, it is just as important to avoid overloading gestures (e.g., reusing the same gesture across applications but mapping it to different outcomes). Complexity can also be related to efficiency, as increased complexity often correlates with decreased efficiency (see Section 3.4.2 for a list of efficiency measures). These factors should thus be considered in tandem.

If a gesture is found to be highly complex, the designer can reduce the number of segments of the gesture movement, the number of fingers or hands required to perform the gesture, or completely redesign the gesture itself.

3.4.2 Efficiency. As per Office, *efficiency* refers to the accuracy and completeness with which users perform an action or gesture [239]. Within the context of gesture design, efficiency is often

⁸See references including [1, 15, 24, 37, 49, 54, 59, 73, 79, 82, 91, 94, 103, 105, 165, 175, 179, 185, 187, 188, 240, 244, 246, 255, 281, 285, 291, 295, 318, 324, 333, 335, 362, 368, 372, 381, 386, 387].

referred to as *difficulty*, *human performance*, *interaction cost*, *speed*, *duration*, *effort*, *easy to perform*, or *easy operation*.⁹ Efficiency is important for gesture-based systems because it enables users to increase their throughput as well as decrease the amount of physical and mental effort that they need to exert. Selecting a single object using a lasso gesture, for example, takes longer, covers a greater distance, and increases the likelihood of incorrect selections when compared to a simple tap gesture. While using a mobile keyboard, for example, typing by tapping letters with one's thumb is less efficient than swiping through the letters in a continuous manner, as the latter removes the "up and down" motion between letter selections [168, 387]. Furthermore, systems which support "lazy" execution, where the general shape of swipes is considered rather than specific locations, enables experts to generate gestural input with increased efficiency [387, 387].

Efficiency is often considered when using Expert-led or Computationally-based methodologies because the experts or algorithms designing the gestures use established techniques such as measuring the time necessary to perform each gesture [278] or by applying estimation models such as Vataavu et al. [330] or Isokoski's [129] efficiency measures, Cao and Zhai's CLC Model or Leiva et al. KeyTime technique for unistroke gestures, Bjerre and Pedersen's [43] or Rice and Lartigue's Touch-level Models [268], or Batran and Dunlop's enhanced Keystroke-Level Model [35], as well as Leiva et al. GATO technique [183].

If the gestures are found to be inefficient, one can make several modifications to the gesture. The duration or size of the gesture could be decreased. A designer could also reduce the number of repetitions that are required to perform the gesture (e.g., change the control to display ratio), or completely redesign the gesture itself. In addition to modifying the gestures, beginning gesture predicting during input can also reduce the time needed for a system to recognize gestures and trigger corresponding actions [41, 286].

3.4.3 Ergonomics. The Office defines *ergonomics* as the human anatomical, anthropometric, physiological, and bio-mechanical characteristics as they relate to physical activity [239]. Thus, with respect to gestures, users should feel physically well when performing a gesture and they should not encounter any pain, tiredness, or discomfort. When designing gestures, it is important to consider the potential risks that may be involved. Some risks may be immediately apparent, such as asking users to perform a complex gesture that requires visual attention on the touchscreen of a car, whereas other may be more long term, such as repetitive movements that lead to repetitive stress injuries or carpal tunnel syndrome [229]. Although many factors encompass ergonomics, it is important a user is aware of the positions that their body will be in when performing a gesture to avoid fatigue, such as with the "gorilla arm" [275]. Within the literature, ergonomics is also often referred to as *physical ergonomics*, *point of articulation*, *safety*, *fatigue*, *comfort*, *stress*, *feasibility*, *bio-mechanical risk*, or *gesture mechanics*.¹⁰

To determine if gestures are ergonomic, designers should evaluate if their proposed gestures will be safe and comfortable to perform. Designers should evaluate each gesture in their vocabulary by examining:

- The use of outer positions of the joints of the finger, wrist, shoulder, and so on. [229]
- The use of repetition [229]
- How relaxed the muscles are [229]
- The use of static positions [229]

⁹Suggested references include [22, 35, 43, 56, 59, 103, 125, 129, 136, 181, 183, 268, 276, 278, 288, 311, 314, 330, 334, 364, 366, 387].

¹⁰See references including [1, 32, 34, 37, 44, 57, 62, 90, 98, 100, 107, 117, 118, 120, 125, 133, 210, 229, 236, 240, 244, 246, 258, 266, 267, 282, 300, 315, 319, 323, 335, 342, 350, 351, 366, 367, 372, 374, 375, 382].

- How relaxed the “middle” neutral position is between outer positions [229]
- The internal and external force exerted on joints [229]
- If the gestures will stop body fluids from flowing [229]
- Increases in heart rate and fatigue [34]
- The amount of arm fatigue [117]
- The prevalence of multi-finger gestures (as opposed to single touch and two-finger gestures) [266]
- The prevalence of bi-manual movements (as opposed to sequential uni-manual movements) [266]
- The flexibility in the number of strokes required for a gesture [266]
- The use of unfamiliar shapes and complex geometries [266]

In addition to these guidelines, a gesture may be safe to perform infrequently but repeated performances of it may fatigue a user or gradually degrade other aspects of their posture (e.g., they begin to slouch, shift their weight onto one leg, lean against another surface, etc.). Thus, it is important to consider the complexity of a gesture, in addition to how it will be recognized, when thinking about its ergonomics. Within the literature, gesture ergonomics are most often associated with Expert-led or Computationally-based methodologies because end users infrequently consider the effects of performing a gesture multiple times over a long duration, or within different, possibly more tiring contexts. There has been some exploration of ergonomics within User-led methodologies, i.e., external physical constraints have been employed to emulate fatigue when eliciting gestures. For example, Ruiz and Vogel fastened weights to users’ wrists to simulate low arm fatigue [282]. However, more exploration is required to identify whether and how accurately external constraints can simulate the many kinds of ergonomic constraints and criteria that were identified from the literature.

Ergonomics also depends on the modality that gestures are performed with. Less physical effort will be exerted if a gesture is performed with fingers than with arms, for example. When directly manipulating elements on interactive surfaces, objects that are too large or too small can be problematic. A common approach to addressing the inherent ergonomic issues of modalities is to require that users perform small or indirect gestures with parts of their body that require the least physical movement [140, 251, 355].

If a gesture is deemed to be unsafe or harmful, a designer may be able to change the size of the gesture, the orientation of the body part making the gesture, where the gesture is performed, the pressure required (if utilized), or the number of fingers required. In severe cases, the gesture itself may need to be completely redesigned [229].

3.4.4 Occlusion. How much of a system’s output is covered or blocked when a gesture is performed is commonly referred to as occlusion or the *fat finger problem* [336]. For example, on a smartphone, if a user wants to place a pin at a location on a map by holding their finger down for a fixed period of time, they would not know if the pin was placed correctly because their finger would occlude the portion of the screen underneath their finger [379]. This problem becomes even more concerning on ultra-small displays such as smartwatches [377]. When content is occluded, the user either misses important information or feedback, has to relocate where they perform a gesture, or change the size, scale, or orientation of the gesture they perform [336]. As such, the occlusion factor is related to the feedback factor and occlusion is of most relevance to systems that require direct input, such as touchscreens or stylus-enabled devices.¹¹

¹¹See references including [27, 31, 40, 50, 51, 74, 93, 201, 223, 279, 291, 336–339, 346, 354, 355, 374, 379, 380, 387].

To determine if a gesture occludes other content, designers first need to consider how the movements each gesture requires will obscure or block any visual information or interaction widgets that may be presented to the user. As proposed by Vogel and Balakrishnan, the Hand Occlusion Model can be used to determine which areas of the screen will be obscured by the hand, wrist, and arm [336]. If this model is not a viable option for other body parts or posture, a designer could also visually inspect the gestures or the intended interface to determine how much content is blocked by the movements of the gestures, similar to Brandl et al. [51]. Because it is possible to determine if a gesture will occlude existing or future content without requiring that a user physically perform the gesture, this factor is often considered during Expert-led or Computationally-based methodologies.

If a designer has determined that one or more gestures will occlude content, a few options are available. First, the designer can maintain the gesture itself but change where it is performed on the screen. Making use of dedicated or offset gesture areas could also be a viable option [50, 381]. Alternatively, one could change the size of the gesture (i.e., instead of requiring motions that span an entire surface, a micro version of the gesture could be required). Call-outs of the occluded content or offsets of the input could also be possible modifications [336, 337].

3.5 System Factors

The final category of factors that emerged were those related to the features and techniques contained within the systems and devices themselves. The two factors within this category arise from the techniques and technologies that are used to *recognize* the user's input and provide *feedback* to the user during and after the gesture has been performed.

3.5.1 Feedback. Another factor that was prevalent in the literature was *feedback*. As defined by Schneiderman, feedback refers to the communication that a system has with a user resulting directly from the user's actions (i.e., gesture [287]). Feedback is important because it allows a user to determine whether the actual functionality of the system matches their mental model of the functionality of the system [233]. Within gestural systems, feedback can inform the user whether a gesture they made is currently being recognized or was recognized. In some systems, this information can be conveyed using a sound [298], whereas in others visual changes will occur within the interface [17] or haptic, olfactory, or tactile feedback may be rendered or generated. Feedback has also been referred to as *feedback ambiguity* or *system response*,¹² and is most often considered when using Expert-led methodologies.

To determine what feedback is appropriate for a given action or functionality, one could elicit feedback mechanisms from users or evaluate the effectiveness of various techniques that were designed by experts during user studies with pre-defined sets of gestures. These studies could measure noticeability, frustration, and so on. The design and evaluation of feedback are similar to the complexity of gesture design itself because there are a multitude of intertwined factors that influence feedback including size, shape, scale, persistence, duration, frequency, modality, and so on [293]. In essence, feedback should clearly communicate success and sources of errors to users without ambiguity so that they can take corrective actions when using a system [356].

Although feedback occurs after the execution of a gesture, feedforward (i.e., the information provided while performing a gesture [36, 75, 76, 275]) is also important to consider as it can improve learnability. For example, with OctoPocus, on-screen feedback, and feedforward are provided while drawing a gesture. The feedback and feedforward enable users to not only understand the system's

¹²See references including [17, 36, 75, 86, 91, 93, 98, 103, 115, 120, 163, 164, 189, 233, 234, 275, 287, 293, 298, 333, 337, 342, 351, 354–356, 379, 384].

status but also be reminded about possible gesture paths, which can improve the execution and learning of the gestures [36].

3.5.2 Recognition. The *recognition* of a gesture is the process of tracking a gesture from its initial representation through to its later conversion into a semantically meaningful command [262]. Within a subsection of the research literature, recognition techniques (i.e., *detectability*, *ease of implementation*, *orientation*, *recognition rate*, *selection accuracy*, *sensing*, or *risk of confusion with natural movements*) have been of the utmost importance.¹³ When a system is unable to recognize a gesture, or recognizes a gesture incorrectly, users will try to perform the gesture again and again in the same way, make larger and more exaggerated movements until it is recognized, or they will stop performing the gesture altogether [40]. Recognition is most often considered when Expert-led or Computationally-based methodologies are used.

Recognition goes hand-in-hand with discoverability, i.e., if a user walks up to a system for the first time and performs the “correct” gesture, the gesture has no utility if the system does not accurately recognize it. Most often, the performance of a recognition technique can be evaluated using existing databases or recordings of gesture performances or in real time by capturing users perform gestures. As there are a multitude of gesture recognition techniques in the literature, the reader is referred to Khan and Ibraheem [150], Murthy and Jadon [222], Suarez and Murphy [307], or Pisharady and Saerbeck [254] for reviews of possible gesture recognition techniques.

3.6 Discussion

The categories of factors were also found to have different design and evaluation requirements. The situational factors, such as context and modality, often need to be approached on a case-by-case basis, where gestures are designed to suit the specific needs of the contexts and leverage the advantages of different modalities. As such, these factors were generally evaluated by examining whether the designed gestures could help users complete tasks under certain contexts and with certain modalities, and depending on the context could be assessed using expert or computational methodologies.

On the other hand, the physical, cognitive, and system factors were found to be general criteria that should apply to all gesture systems that are designed. Because of this, physical factors and system factors were generally evaluated using datasets or real-time motions that were collected from users but were most often only considered during Expert-led or Computationally-based methodologies. Cognitive factors were generally evaluated via questionnaires with participants during or at the conclusion of an experiment to capture their impressions but not impede with their performance of the gestures. These factors were often assessed only during User-led methodologies.

Thus, as the factor identification found, there are many factors that are implicated in the design of gestures (Table 2). Some factors stand relatively isolated, such as social acceptability, whereas others have a high interdependence, such as learnability and discoverability. Across the literature that was reviewed, there was not one experiment or methodology that considered all of these factors simultaneously, thereby underscoring the importance of reviewing and synthesizing this body of literature and identifying these factors. In what follows, we assess the three categories of gesture design methodologies in more detail, focusing on the implications that can occur when one chooses to use a given methodology to design a gesture vocabulary.

¹³Suggested references include [12, 20, 24, 29, 34, 40, 64, 88, 90, 100, 108, 109, 119, 124, 132, 136, 141, 147, 150, 151, 158, 160, 163–167, 172, 186, 187, 195, 203, 207, 219, 222, 230, 234, 238, 241, 244, 254, 257, 262, 264–266, 270, 273, 278, 283, 292, 300, 307, 314, 315, 321, 327, 342, 343, 348, 351, 354, 372, 387, 389].

Table 2. A Summary of the 13 Factors Identified within the Literature Review as being Important to the Design of Gesture Vocabularies

Factor	Definition	Evaluation	Refinement
Context	Information that characterizes the situation where a gesture will be performed (i.e., people, places, and devices) [79]	N/A (pre-condition)	Understand requirement by working through: <ul style="list-style-type: none"> - Scenarios [308] - Storyboards [313] - Use cases [131] - Task analysis [81] - Wireframing - Volere Requirements Specification Method [272]
Modality	The sensory input channel used to perform a gesture and the output channel the system uses to provide feedback to the user [72]	N/A (pre-condition)	Input Modalities: <ul style="list-style-type: none"> - fingers [265, 302, 303, 315, 374] - hands [187, 241, 244, 267] - feet [134] - nose [385] - entire body [69] - styli [63, 316] - controllers [64, 372]
Social Acceptability	The social and cultural factors that affect the user experience when a gesture is performed [358]	<ul style="list-style-type: none"> - Visually inspect for weirdness or attention seeking behavior [269], interference with communication [269], and uncommon movements [269] - Ask users during / after user study, after in the wild deployment, or in online survey 	<ul style="list-style-type: none"> - Make gestures small, unobtrusive, unnoticeable, similar to everyday movement [211, 289] - Make gestures physically comfortable and enjoyable [289]
Discoverability	The quality of gestures that enables a user to access intended referents of those gestures despite a lack of knowledge about the gesture [361]	<ul style="list-style-type: none"> - Compute agreement score [362] - In the wild deployment and observation [289] 	<ul style="list-style-type: none"> - Utilize adopted gestures - Mimic common actions [121] - Reveal gestures when appropriate [18]
Intuitiveness	The degree to which a gesture makes use of transferable and existing knowledge [260]	<ul style="list-style-type: none"> - Lab study with matching gestures to commands [228, 362] and matching commands to gestures [228] - Lab study to measure differences between expert and novice duration [3] 	<ul style="list-style-type: none"> - Leverage existing and transferable knowledge

(Continued)

Table 2. Continued

Factor	Definition	Evaluation	Refinement
Learnability	The ease with which new users can begin effective interaction and achieve maximal performance with a gesture [83]	<ul style="list-style-type: none"> - Lab study to evaluate how many gestures can be remembered [21, 34, 50] - Compute the similarity of gestures [194] - Compute number of segments in a gesture [364] - Compute duration of gesture performance 	<ul style="list-style-type: none"> - Reduce the number of gestures - Reduce the number of gestures required a unit task [364] - Reduce the similarity of gestures [194] - Incorporate intuitive mappings - Incorporate gesture training systems [13, 36, 50, 381]
Transferability	The degree to which the learning of a gesture can enhance or undermine its performance in another context [248]	<ul style="list-style-type: none"> - Lab experiment with retention and transfer paradigm [13, 17] - Lab study asking users to transfer learned gesture to another context [294, 295] 	<ul style="list-style-type: none"> - Modify gestures that can be chunked [56] - Incorporate gesture training systems [13, 36, 50, 381]
Complexity	The state or quality of a gesture being intricate or complicated [243]	<ul style="list-style-type: none"> - Inspect number of steps or movements required by gestures [285] - Inspect movements that require substantial effort [381] - Indicated by the measurement or predication of production time 	<ul style="list-style-type: none"> - Reduce the number of segments in gesture movement - Reduce the number of figures and hands required - Match the complexity of the gesture to that of the task
Efficiency	The accuracy and completeness with which users perform an action or gesture [239]	<ul style="list-style-type: none"> - Compute production time [278] - Apply computational and estimation models for pen input[59, 129, 321], and touch input [35, 268] 	<ul style="list-style-type: none"> - Reduce the size of gesture - Reduce the duration of gestures (i.e., shorter dwell time) - Reduce the number of repetitions
Ergonomics	The human anatomical, anthropometric, physiological and biomechanical characteristics as they relate to physical activity (i.e., performing a gesture) [239]	<p>Look for:</p> <ul style="list-style-type: none"> - Use of outer positions [229] - Use of repetition [229] - Use of static positions [229] - How relaxed the muscles are [229] - How relaxed the “middle” neutral position is between outer positions [229] 	<ul style="list-style-type: none"> - Change size of the gesture - Change where the gesture is performed - Change the geometry, pressure of the gesture to avoid issues described on the left

(Continued)














Table 2. Continued





Factor	Definition	Evaluation	Refinement
		<ul style="list-style-type: none"> - The internal and external force exerted on joints [229] - Whether gestures will stop body fluids from flowing [229] - Prevalence of multi-finger gestures (v.s. single touch and two-finger gestures) [266] - Prevalence of bimanual movements (v.s. sequential unimanual movements) [266] - Use of unfamiliar shapes and complex geometries [266] Also: <ul style="list-style-type: none"> - Compute the flexibility in the number of strokes required [266] - Determine if they increase heart rate and fatigue [34] - Assess the amount of arm fatigue [117] 	
Occlusion	How much of a system's output is covered or blocked when a gesture is performed [336]	<ul style="list-style-type: none"> - Use Hand Occlusion Model [336] - Visually inspect for occlusion [51] 	<ul style="list-style-type: none"> - Change size of the gesture - Perform at dedicated or offset area [50, 381] - Use callout of the occluded content [336, 337]
Feedback	communication that a system has with a user resulting directly from the user's actions (i.e., gesture) [287]	<ul style="list-style-type: none"> - Lab study with think aloud protocol or post-experiment questionnaire - In the wild deployment with post-deployment feedback 	<ul style="list-style-type: none"> - Select appropriate modality - Indicate success or sources of errors clearly [356]
Recognition	Process of tracking a gesture from its initial representation through to its later conversion into a semantically meaningful command [262]	<ul style="list-style-type: none"> - Compute accuracy/error 	<ul style="list-style-type: none"> - Reduce similarity of gestures - Utilize alternative recognition algorithms [150, 222, 254, 307]

4 HOLISTIC IMPORTANCE OF THE IDENTIFIED FACTORS TO GESTURE DESIGN

The identification and synthesis of the 13 aforementioned factors should provide the community with a more consistent vocabulary and understanding of the characteristics that are important to consider when designing a gesture vocabulary. The analysis also provides an opportunity to holistically explore how these factors influence the gesture design process that one chooses to employ and the implications that such factors can have if they are not considered when designing a gesture vocabulary. Herein, we re-examine the three categories of gesture-design methodologies (i.e., Expert-Led, User-Led, and Computationally-Based) from a factor-oriented perspective to better understand the degree to which each methodology focuses on each factor and highlight the poten-

Table 3. A Comparison of the Degree to which Expert-Led, User-Led, and Computationally-Based Gesture Design Methodologies Focus on the 13 Factors Identified by the Literature Review

Factor	Expert-Led	User-Led	Computationally-Based
 Context	◐	●	○
 Modality	●	○	◐
 Social Acceptability	◐	●	○
 Discoverability	○	◐	○
 Intuitiveness	○	◐	○
 Learnability	○	◐	○
 Transferability	●	○	○
 Complexity	●	◐	●
 Efficiency	●	◐	◐
 Ergonomics	●	○	◐
 Occlusion	●	○	◐
 Feedback	●	○	◐
 Recognition	●	○	●

 – Situational factors;  – Cognitive factors;  – Physical factors;  – System factors; ● – Yes; ◐ – Maybe; ○ – No

tial pitfalls of using such methods by applying the definitions and heuristics synthesized from the literature (Table 3).

4.1 Situational Factors: Context

Among the three methodologies, User-led methodologies best address the context of gesture vocabularies because end-users understand the context within which they will be performing gestures so they are often the best entities to relate designed gestures to targeted tasks and environments [179].

During Expert-led methodologies, if an expert is not well-versed or does not have extensive experience with the context within which gestures will be performed, then it is important that such methodologies contain a phase or component that involve the target end-user population (or a reasonably close approximation of them) in the design process. Such a phase will enable the expert to compensate for the knowledge they are missing and gather much needed insights and feedback from users. For example, when designing gestures for users with visual or motor impairments, it is essential to involve the target population during the design process so that the requirements and preferences of such users can be captured [142, 253].

Computationally-based methodologies are the least effective at understanding and designing for the nuances of context because it is difficult for algorithms to predict or quantitatively measure characteristics such as who will be using a gesture vocabulary, where they will be doing so, and what devices they will be using. In some sense, Computationally-based tools such as MAGIC [26] and MAGIC 2.0 [160] could be viewed as tools that understood a primitive level of context because they sought to learn and identify motions that should not be recognized (e.g., false positives or unintentional input within everyday use), however, this level of context is far more general and vague than that which could be obtained when User-led or Expert-led methodologies are used. If,

however, a researcher explicitly provides such information to a computational model, such as how Stern et al. specified task-based information as part of their multi-objective optimization algorithm [300], then it is possible for such methodologies to consider context during gesture design. Such methodologies, however, do require that the expert, programmer, or designer has knowledge about the context to provide to such algorithms.

4.2 Situational Factors: Modality

Expert-led methodologies excel at considering the implications of different modalities because experts are much more familiar with the characteristics and limitations of a chosen modality than users, especially for emerging technologies that they may have created [18, 119, 376, 377]. For example, by instrumenting a digital stylus with multi-touch sensing capabilities along the entire barrel of the stylus, Hinckley et al. were able to recognize and identify a new set of stylus barrel gripping gestures that could be used to enhance interaction with multi-touch surfaces [119].

Computationally-based methodologies could be used to evaluate the influence of a modality on a gesture vocabulary, but only if the properties and functionality of a modality can be modeled or recorded by experts. For example, with the aCAPella [106], Exemplar [111], and Mogeste [245] systems, in addition to work by Kim and Nam and Kim et al., designers or developers could record data from modalities such as touchscreens, mobile phones, cameras, RFID tags, or sensor-embedded objects and use these data streams to generate gesture vocabularies [152, 153]. These approaches, however, require that the designer or developer has background knowledge and insights into which data they should record from such devices and has the ability to access and record such data.

Unlike experts, who's knowledge about a modality comes from their authoring of the modality or their experience watching participants use the modality, users are far less likely to be experts with a given modality. Interestingly, many research projects have solicited gestures from novice users who were using a modality or device for the first time (e.g., interacting with smart glasses [317] or unmanned aerial vehicles [249]). While novices' limited experiences may help them create gestures that are more discoverable, their naivete may encourage them to reuse gestures from existing applications or contexts [85, 214, 281] or design gestures that would be awkward or socially unacceptable to perform [204]. This thus makes User-led methodologies the least effective at evaluating the role of modality.

4.3 Situational Factors: Social Acceptability

The best practice to ensure gestures are socially acceptable is to involve end users in the design process and to encourage them to actively think about the social context that the gestures will be performed in [2, 82, 99, 269]. User-led methodologies are thus the best methodology to use to do so.

By default, Expert-led methodologies, are not the preferred method to solicit the social acceptability of gestures because, unless the expert(s) are experts in the nuances of the social context their gesture vocabulary will be used, they cannot understand the social or cultural implications of each gestures' performance. However, if such methodologies are augmented to allow participants to provide feedback using think-aloud protocols or questionnaires, then such methodologies can be used to understand how participants feel while or after performing gestures [2, 5, 82, 99]. For example, Expert-led gesture vocabularies that were created without considering social acceptability recommended that interaction with remote displays should occur at substantial distances from a display (e.g., 250 cm) [205, 275]. However, research by Ahlström et al., found that in-air gestures performed further than 30 cm away from a device or for longer than 6 seconds were perceived as significantly less socially acceptable in public settings than those performed closer to a device or for a shorter duration [5].

Computationally-based methodologies are, unfortunately, not suitable to ensure gestures are socially acceptable because the many variables that are relevant to social acceptability, such as culture, audience, and environment cannot be easily quantified and often vary immensely from situation to situation.

4.4 Cognitive Factors: Discoverability

Discoverable gestures are often informed by the past knowledge, experiences, and habits of users [361]. In many cases, the history that a user brings to an elicitation study differs significantly from the history and experiences of experts and is difficult to capture within a computational model. As such, it is often difficult for experts or computational models to conceive gestures that will be discoverable by general or target populations. For example, both Hinckley et al. and Xia et al. introduced expressive bimanual pen and touch interaction techniques. However, participants reported that they would not have “guessed” many of the gestures without guidance [121, 376].

User-led methodologies, however, seek to mimic users’ first attempts at a gesture in the wild with the hope that the derived gestures will be easily discoverable. However, individual differences between users make it challenging for elicited gestures to be universally discoverable. For example, Wobbrock et al. found that among the 27 commands for which participants created gestures in their study, only 2 gestures (i.e., *move a little* and *move a lot*) were universally agreed upon by all participants [362]. As suggested by Vatavu and Wobbrock, in some cases, one may need to compute how dissimilar the gestures for a given referent are [332]. While gestures that receive moderate agreement ratings may still be discoverable after several attempts, the lack of consensus between users suggests that gestures that are created via User-led methodologies can improve discoverability but do not fully guarantee it.

4.5 Cognitive Factors: Intuitiveness

Similar to discoverability, intuitive gestures also rely on users’ past experience and first hand knowledge. Gestures designed by end users should be intuitive, however, Wobbrock et al. found low agreement amongst users for the majority of the referent commands in their study [362]. In addition, while gestures such as *double-tap*, *touch and hold*, *pinch* and *spread* were all included in the UDG set in Wobbrock et al. work [362], these same gestures were not discovered or understood by the older adult participants in Leitao and Silva’s work [179]. As such, gestures that are created via User-led methodologies may have increased intuitiveness but may not necessarily be intuitive.

In a similar vein, the use of Expert-led methodologies does not guarantee that one will solicit an intuitive gesture vocabulary, as experts’ prior experience and (likely) overexposure to gestural interaction may bias what they consider to be intuitive [224]. Lastly, because it is difficult for computational models to predict or create meaningful cognitive mappings, the use of Computationally-based methodologies is also not suitable for the creation of intuitive gesture vocabularies.

4.6 Cognitive Factors: Learnability

Gestures that are developed via User-led methodologies build upon users’ collective past experiences and knowledge, and are selected based on how high of an agreement score all of the gestures that were elicited for a given referent are [9, 58, 326, 362]. In theory, this should result in gestures that are easy to learn because they are the most common gestures that were elicited from participants, and are thus the most familiar or easy to recall. However, because participants rarely get to reflect or refine the entire set of gestures they have created during an experiment [82, 84, 105, 295, 370, 373], they do not have the opportunity to holistically consider how similar subsets of gestures may be, and thus how easy it would be to learn the entire gesture vocabulary.

If participants are provided with refinement opportunities, then User-led methodologies could be a viable method to evaluate the learnability of a gesture vocabulary.

In terms of Expert-led methodologies, if an expert can assess how similar a subset of gestures is (e.g., [19, 197, 330]) or is aware of best practises for gesture design (e.g., use mnemonics [277] semaphoric or pantomimic mappings [6], be analogous to the physical effects of the real world [387], and so on) avoid, then it is possible that experts could perform preliminary assessments about the learnability of a gesture vocabulary. However, because experts have extensive experience watching users perform gestures or interact with devices, and have immense experience interacting with technology themselves, their very familiarity with gestural interaction may lead them to underestimate the difficulty that a target population of users may have while learning a gesture vocabulary. This was demonstrated by Nacenta et al., who found that gesture vocabularies that were created via User-led methodologies were easier for participants to recall after 24 hours and were preferred to gesture vocabularies that were created by experts [224].

As the literature does not yet have a clear understanding of the role that discoverability and intuitiveness have on the learnability of gesture vocabularies, and there is much ongoing research attempting to understand how to best teach gestures to users (e.g., [13, 36, 49]), it is difficult to empirically model or abstract which movement and function mappings will be best for a vocabulary of gestures. As such, Computationally-based methodologies do not currently support the evaluation of gesture learnability.

4.7 Cognitive Factors: Transferability

Transferable gestures are those that can be executed using different modalities, in different contexts, or with different devices [13, 17, 369]. To increase the likelihood of transferability, some experts have suggested that chaining or chunking techniques may be useful as they allow constituent gesture components to be reassembled as needed [171, 325], while others have recommended the use of mindful abstraction and performance metaphors during training [369]. Because the design of such gestures requires substantial expertise and holistic knowledge about user behaviours and a system's functionality, and also require a deep understanding of the characteristics of different modalities, contexts, and devices [325], the use of Expert-led methodologies that involve end users in the evaluation process [13, 17] are the only viable way to design gesture vocabularies that fulfill such needs.

Although users may reuse gestures that they are familiar with during User-led methodologies, this does not guarantee that the gestures they propose will be appropriate for other devices, contexts, or applications [214]. Because understanding how transferable gestures are required expertise [325], it is unlikely that users will be able to design gestures that are universal or applicable to a variety of contexts. As such, User-led methodologies are an ineffective methodology to use to ensure that gestures are transferable.

As the research literature does not yet have a comprehensive understanding of the characteristics of gestures that make them transferable, there are no models or frameworks that can be integrated within a computational-based gesture design methodology to evaluate the transferability of a gesture vocabulary. Thus, similar to User-led methodologies, computational-based methodologies are not suitable to assess gesture transferability.

4.8 Physical Factors: Complexity

As noted in Section 3.4.1, there are many measures that can be computed by experts or integrated within computational models to measure the complexity of a gesture, such as counting the number of steps that need to be performed, the number of discrete movements that are made, or the number of repetitive movements that are required [285]. Experts or computational models can

also make use of more complex measures such as those found in Leiva et al. Omnis Prædictio software, which can measure the turning angle, density, aspect, and perimeter to area ratio of a gesture [185]. Because these measures exist, and they are trivial to compute, both Expert-led and Computationally-based methodologies are suitable methodologies to ensure that the complexity of individual gestures, and entire gesture vocabularies themselves, are as low as possible.

On the other hand, with User-led methodologies, it is unlikely that users will create gestures that are complex to produce or intricate to perform, especially as users are likely to reuse gestures from other devices or situations during elicitation studies [214]. However, because not all User-led methodologies provide users with the opportunity to review the entire set of gestures they have proposed or see the final gesture set after acceptability criteria have been applied [82, 105, 295], they may neglect to consider the overall complexity of an entire gesture vocabulary. For this reason, depending on the specific instantiation of User-led methodology that is utilized, users may or may not be able to evaluate the complexity of the gestures that they create.

4.9 Physical Factors: Efficiency

Expert-led methodologies are a great way to evaluate efficiency because experts can harness established techniques or metrics within their experimental software to measure how efficient a gesture is. The simplest such metric is the time necessary to perform each gesture [278], however, depending on the type of gesture and input modality, one could also apply Vatavu et al. [330] or Isokoski's [129] efficiency measures, Cao and Zhai's CLC Model [59], Leiva et al. KeyTime technique for unistroke gestures [181], Bjerre and Pedersen's [43] or Rice and Lartigue's Touch-level Models, Rice and Lartigue [268] or Batran and Dunlop's enhanced Keystroke-Level Models [35], or Leiva et al. GATO technique [183].

Because the models and techniques that experts can use within Expert-led methodologies utilize much of the same type of information as that required by Computationally-based methodologies, one could imagine that such efficiency measurement techniques could be integrated into the software used for Computationally-based gesture design. While such methodologies currently do not include measures of efficiency, they would allow for such computations to occur automatically alongside evaluations of recognizer accuracy, thus further increasing the utility of Computationally-based methodologies.

Because the focus of many User-led methodologies is to find mappings between functions and movements that are easy discover, it is not common for users to be concerned with efficiency [334]. As users are often asked to design one gesture per referent or perform a small number of repetitions of a movement until they are "happy" with the gesture they have made [334, 340], users thus have limited opportunities to become familiar with the gestures they have designed or perform them for extended periods of time. This often results in User-led methodologies being unable to evaluate efficiency unless users are explicitly asked about how easy the gestures they have designed would be to perform [22, 276, 282, 362].

4.10 Physical Factors: Ergonomics

As identified within research by Barclay et al. [34], Hinckley et al. [118], Nielsen et al. [229], and Rekik et al. [266], there are *many* facets of ergonomics that one should attend to while creating gesture vocabularies. Although users may have personal experience with the effects of performing repetitive or tiring movements [221, 225], unless they are explicitly told to attend to characteristics such as the forces exerted on the joints while performing various movements, which types of movements can stop body fluids from flowing, or the arm fatigue that may occur when performing certain motions [282], it is unlikely that they will consider these characteristics. As such, User-led

methodologies that do not inform the user about these characteristics are the least suitable to use when creating gesture vocabularies.

On the other hand, many experts are aware of the role that such ergonomic characteristics can have during gesture production. In some cases, this may be due to their years of experience watching participants complete tasks during user studies (i.e., seeing the gorilla arm phenomenon occur first hand [44]), receiving feedback about how uncomfortable certain gestures are to perform [7, 80, 151, 165, 269], or their awareness of the literature on gesture ergonomics [57, 107, 210, 246, 258]. Regardless of how they obtained their expertise about ergonomics, the use of Expert-Led methodologies allows these expert designers to implicitly evaluate their gesture vocabulary for ergonomics issues throughout the implementation, testing, and debugging phases of their experimental software or gesture recognition algorithm development. This unique opportunity to iteratively refine a vocabulary through first-hand experience cannot be found with other types of methodologies.

Much work has also been conducted to model the bio-mechanical characteristics of the human body, and has resulted in a number of bio-mechanical models or metrics that describe movement. For example, work by Hincapié-Ramos et al. created the consumed endurance metric, which is the ratio of total time to endurance and can be used to characterize the gorilla arm effect during mid-air interaction [117]. In similar research by Jang et al., the cumulative arm fatigue of mid-air interaction was computed using calculations about a user's perceived exertion and arm motion kinetics [133]. Other research such as that by Bachynskiy et al. identified 11 clusters of bio-mechanical movements that had distinct performance, muscular, and spatio-temporal characteristics and could be used to summarize muscle movement during interactive tasks [32]. Although such models and metrics have yet to be integrated into the computational-based gesture design tools that are available today, one could assume that such integration would enable the automatic detection of some characteristics that may be too complex or dangerous for a human designer to test and evaluate themselves, e.g., if a gesture stopped body fluid from flowing to a certain body part or after how many repetitions a repetitive stress injury may occur.

4.11 Physical Factors: Occlusion

Because experts are aware of the graphics or information that will be provided on an interface before, during, and after the production of a gesture, it is unsurprising that experts, and Expert-led methodologies, are best suited to consider issues related to occlusion. As demonstrated by Brandl et al. work with multi-touch tabletops and styli, experts often have ample prior experience and knowledge about how the hands or arms may move around a surface and could block or occlude content that is displayed [51]. In the case of Brandl et al. research, the researchers used their prior knowledge to create handedness-aware menus and adapt the location of UI widgets so that the information they presented to users would not be covered by users' hands or styli.

Although not currently integrated within the computational tools that have been designed to aid in the development of gesture vocabularies, computational models such as Vogel et al. model of hand occlusion would be a very useful addition [339]. As this research showed, the small variety of hand, wrist, and forearm postures that are used while inking can be abstracted into a circle and rectangle unit that can be reoriented around the position of a stylus nib, provided that the location of the stylus nib and the handedness of the user is known. This model can then be used to relocate the position and orientation of graphics and UI widgets on an interface [336]. Similar models have also been developed for single user [74, 338] and multi-user [223] touch-based interaction on multi-touch tabletops [27, 279, 380].

User-led methodologies are the least effective at considering occlusion because users are often unaware of how their, or others', movements can occlude information that is being presented [31].

In addition, these methodologies typically use a simple test environment to elicit gestures or ask participants to make modifications to their movements to overcome technical limitations, thus making it difficult for users to fully understand the visual context that will be present during the real world performance of the gestures [62, 135, 157, 255, 319, 321, 333, 362]. These practises, and thus the use of User-led methodologies, decrease users' abilities to design gestures that will have a negative influence on user experiences.

4.12 System Factors: Feedback

The design of effective feedback or feedforward solutions for gestural systems requires one to have a holistic understanding of the ways in which the timing, duration, modality, semantic meaning, and presentation schedules of feedback will influence gesture vocabularies [13, 36, 293]. Given that much expertise is required to (i) be aware of such factors and their interdependencies, (ii) understand how to measure them, (iii) be able to design feedback techniques that fulfill them, and (iv) be able to interpret users' reactions to them, it is unsurprising that Expert-led methodologies are the only viable methodologies to use when designing effective feedback for gesture vocabularies.

While Computationally-based methodologies may be able to help designers and developers assess how the timing or duration of a gesture can influence its recognition rate, current tools and software are unable to simulate users' responses to various forms of feedback. If such tools were extended to include information from crowdsourced studies that evaluated different feedback techniques, similar to how CrowdLearner enabled the rapid creation of mobile gesture recognizers [11] and the crowd was used to evaluate the social acceptability of spatial user interactions on head-mounted displays [8], then it may be possible for Computationally-based methodologies to be viable solutions to evaluate feedback in the future.

Although users may be quick to recognize when they perform an action and do not receive any feedback [36, 75, 93, 164, 189], their lack of knowledge about how an interface senses their input, how it determines if feedback should be provided, how it determines which feedback should be provided, and how it generates or visualizes feedback results in users, and thus User-led methodologies, being unable to evaluate feedback during gesture design.

4.13 System Factors: Recognition

Computationally-based methodologies are the best suited of the three categories of methodologies to ensure that gesture vocabularies can be correctly recognized by a system because almost all of the tools that have been created to assist with gesture design include recognition accuracy as a measure that can be quantified or visualized [88, 158, 172, 195, 207, 278, 292, 327]. Because designers and developers can directly see potential recognition performance in the same tool that they are using to generate gestures, it becomes trivial to use such methodologies to evaluate the impact that different recognition algorithms, datasets, or sensing techniques have on a gesture vocabulary.

While creating gestural interfaces and running experiments to evaluate them, it is common for experts to have some form of gesture recognition within their system, either via a series of algorithms that monitor sensor data (e.g., [36, 48, 63, 224, 377]), or by using a Wizard-of-Oz approach (e.g., [149, 154, 297]). Thus, many experts are aware of which data would be needed to recognize a gesture, how the similarity between gestures can influence recognition, and the role that poor recognition would have on user experiences. Therefore, it is not uncommon for Expert-led methodologies to consider recognition during gesture vocabulary creation.

Because many of the users who participate in experiments that are conducted using User-led methodologies are not developers or designers, but rather laypeople, their knowledge about how a technology works, the data that can or is sensed while a gesture is performed, or the techniques

that are available to recognize a gesture, is lacking [71, 105, 135, 258, 334]. Thus, they are often the least qualified to make judgements about how recognizable a single gesture or an entire gesture vocabulary will be, resulting in User-led methodologies being the least suitable methodology to evaluate this factor.

4.14 Summary

The above evaluation used a factor-centric view of the three categories of gesture design methodologies that are commonly used within the literature. Based on this evaluation, it appears that User-led or Expert-led methodologies are better suited to consider situational factors, however, the end users involved in these design processes may not have the prior knowledge required to leverage the strength of a given modality. User-led methodologies also seem suitable to evaluate some cognitive factors, however, novice users may not be aware of the many issues that can hamper these factors, so it is difficult to definitively recommend a category of methodologies to use. Expert-led and Computationally-based methodologies appear to excel at supporting physical and system factors because the extensive knowledge and expertise that experts have accumulated, and the existence of models derived from this knowledge, can be integrated into systems to evaluate these types of factor categories.

Although each methodology has its own strengths and weaknesses, the analysis demonstrated that no one methodology is currently capable of considering all factors, however, Expert-led methodologies are capable of evaluating a subset of factors from each factor category. As it is currently unclear the degree to which these methodologies could be modified to consider each factor or more factors within a factor category, we next identify research directions that are needed to clarify and identify how these methodologies could be used more fruitfully in the future.

5 FUTURE RESEARCH AGENDAS

The review of the literature on gesture design, the identification of the factors that are important to consider when designing gesture vocabularies, and the analysis of how different methodologies do or do not support the evaluation of different factors, suggested that there are many research avenues that should be considered to improve the design of gestures. In what follows, we present a potential gesture design process that could be used to iteratively evaluate a subset of factors that a designer deems appropriate for a given use case. We also discuss some additional factors that arose during the literature review and highlight how they could impact the ever-changing ecosystem of gesture design.

5.1 A Potential User-Centered, Factor-Centric Gesture Design Process

As indicated above, one limitation of existing gesture design approaches is that there is a mismatch between the sets of factors a methodology is suitable for and the sets of factors gesture designers may wish to optimize. Rather than selecting one methodology and being unable to evaluate a given factor, one possible alternative is to take a user-centered, factor-centric approach, wherein one uses a process that enables different factors to be iteratively refined, evaluated, or removed during design time, whilst respecting the inter-dependencies that exist between various factors.

One possible instantiation of such a design process is depicted in Figure 2. Within this example process, there are four phases that a designer would work through while designing their gesture vocabulary: identifying the preconditions that will dictate the user experience with the gestures, assigning prioritization weights to each of the evaluation factors based on the requirements from the preconditions, creating an initial gesture design using an existing methodology (or using a stock gesture vocabulary), continually refining and evaluating the vocabulary keeping in mind the weights that were assigned to different factors, and finally, arriving at the final gesture vocabulary.

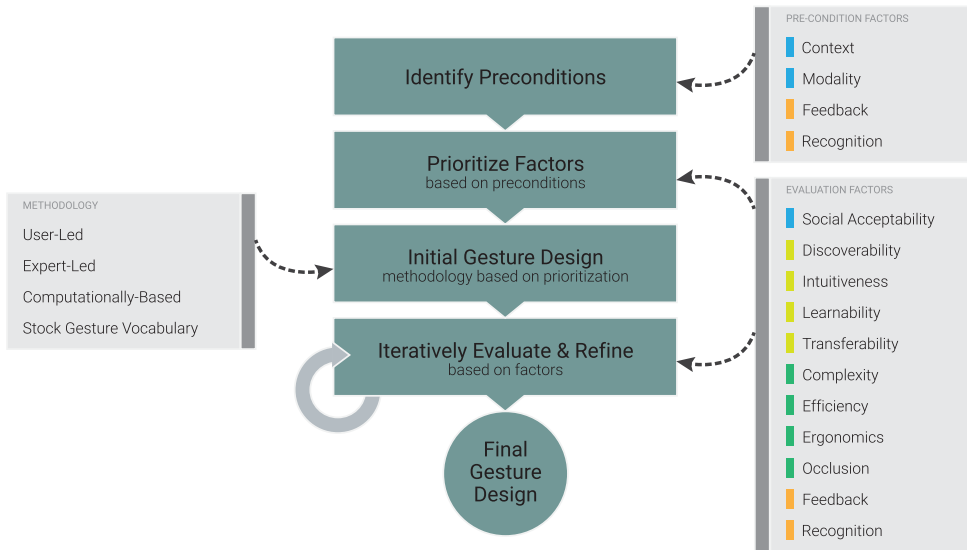


Fig. 2. A proposed iterative, factor-centric gesture design methodology that ensures that the resultant gesture vocabulary satisfies the requirements of a designer’s preconditions and other high-priority factors.

5.1.1 *Understanding Requirements and Prioritizing Factors.* With such an approach (Figure 2), designers would first determine and understand the requirements of their gesture vocabulary. In this example, user experience factors such as context, modality, recognition, and feedback, which will influence the user experience and the designer’s goals, are examined to determine the degree to which other factors should be weighed when iterating on a gesture vocabulary. For example, if a system is to be used with a mall kiosk, the discoverability of the gestures might be weighed higher than ergonomics because a given gesture may be performed infrequently. The identification of requirements will not only help designers form a complete understanding of the goals that are guiding their user experience, but also serve as a sanity check to ensure that a designer is aware of all the challenges their gesture vocabulary may encounter.

As many have demonstrated, there are tradeoffs between many of the factors that were identified in [224, 266, 330, 335]. As these tradeoffs make it impossible to design a gesture vocabulary that would fulfill all the requirements of the factors, a designer should then prioritize the evaluation factors (and reconsider the effectiveness of the feedback and the recognition accuracy), given the requirements of the preconditions. One way to do this would be for a designer to assign each factor an importance descriptor, such as *high*, *medium*, or *low* [353]. As an example, if the end user population is blind, then it may be appropriate to rate factors such as Occlusion or Social Acceptability low, whereas factors such as Discoverability and Recognition should be rated high.

It is also worth noting that executing the various evaluation and refinement methodologies also requires time and resources (e.g., equipment and algorithms). Therefore, we recommend gesture designers consider the accessibility and suitability of the different methodologies regarding each of the factors when adopting this new gesture design approach.

5.1.2 *Iteratively Evaluating and Refining Gesture Vocabulary.* After prioritizing the factors, a designer could then use their preferred method of choice to seed an initial gesture vocabulary (e.g., using a User-led, Expert-led, Computationally-based methodology, or using an existing, stock gesture vocabulary). Once the initial gesture vocabulary has been developed, the designer could then

work through each of the factors based on their priority, iteratively evaluating, refining, and modifying the proposed gesture vocabulary using the methods summarized in Section 3 and Table 2. When a tradeoff must occur, factors with high or medium ratings should receive priority over lower rated ones. Factors that are interrelated, such as Complexity and Efficiency, should be evaluated holistically so that any changes that result do not cause a cascading effect in subsequent iterations over the gesture vocabulary.

The point at which each factor is considered should also be determined based on how likely each factor would be to cause detrimental ripple effects across the entire gesture vocabulary, similar to what is recommended for factors that are interrelated. Factors that, when considered, would require modifications to a number of gestures or the entire vocabulary, should have higher priority and thus be evaluated before those that would require minimal modifications. For example, a factor such as Social Acceptability could require that an entirely new gesture be created to conform to cultural or location norms. In comparison, there are user interface changes that could be implemented to decrease the Occlusion of a gesture. Therefore, within the context of this example, Occlusion should be considered later in the process than Social Acceptability.

A challenge of any iterative design process is to determine when to stop. Most often, an endpoint is reached when one runs out of time or resources, however, achieving the perfect gesture vocabulary may never be possible as the design process itself is inherently one large set of tradeoffs. After iterative evaluation and refinement, the resulting gesture vocabulary should fulfill the requirements of the preconditions and the factors that were deemed to be high importance because important factors should have received enough revision, and been revised early enough, during the process.

5.1.3 Summary. Although this process is but an example, we believe that the creation and use of user-centered, factor-centric approaches have the potential to ensure that the gesture vocabularies that are designed will be the most usable, safe, and beneficial for their target populations. Different from existing gesture design methodologies, which focus on designing suitable gestures by leveraging the expertise, experience, and capability of experts, users, or systems individually, taking a user-centered, factor-centric approach to gesture design should enable the benefits of each different methodology to be used, when appropriate.

Future work is obviously needed to validate this new user-centered, factor-centric gesture design approach by comparing it with existing methodologies, not only to ensure that the resulting gesture approach results in more appropriate gesture sets, but also to evaluate the costs associated with their use (e.g., number of iterative cycles needed, monetary cost of running multiple user studies, etc.) and gain insights into reducing the cost of iterations. If for nothing else, we hope that the proposal and outline of such an approach will improve awareness about the challenges inherent in the design of gestures, the dependencies that exist between different factors, and the potential pitfalls of using the most popular or easy-to-use method to create gesture vocabularies.

5.2 The Evolving Ecosystem of Gesture Design Factors

Although the literature review and analysis identified 13 factors that are critical for the design of gestures, it may never be possible to identify and distill the complete set of factors that influence how users remember, learn, and perform gestures, and how systems can support them in these tasks. We first discuss the role of factor interdependencies on gesture design, detail some additional, higher-level factors that emerged, and then provide commentary on the “naturalness” of gestures.

Understanding the complex connections among various factors is essential for gesture design, as it enables designers to better understand the implications of prioritizing different factors and resolving tradeoffs when designing gesture vocabularies. The identification of the 13 factors was

a first step toward constructing a comprehensive picture of gestural interaction. However, much future work is needed to better understand how to disentangle factors such as Complexity and Efficiency, Intuitiveness and Learnability, Ergonomics and Recognition, Occlusion and Feedback, and so on, as well as if and how they should be designed for independently. In addition to the identified factors, it is also possible that new primitive factors may arise and should be added to the ecosystem of factors as the field furthers its understanding of gesture design. While the user-focused, factor-centric approach that was proposed is extensible and should allow for the addition of new factors, care must be taken to identify and understand the relationships that exist between the newly discovered and existing factors. To encourage the inclusion of new factors into the ecosystem, we plan to construct an online, accessible, and community-maintained database and visualizations of the factors and relevant research articles about these factors to ensure the ecosystem can stay relevant and will evolve with time.

During the literature review, we encountered three other factors that are not included in the final set of 13 factors. The first factor that emerged was the accessibility of gestures, e.g., for older adults [179, 253], as well as with users who have motor [388], or visual impairments [142, 143, 145]. Within the factors that we identified, we did not include accessibility as a distinct factor, because accessibility is a higher-level factor that can be decomposed into a number of the factors that we identified, such as Context, Complexity, Learnability, and Feedback. For example, accessibility is related to situational factors such as Context, because the requirements of users' accessibility need to be understood and considered by designers. It is also related to cognitive and physical factors such as Complexity and Learnability, because a user's cognitive and physical abilities may pose different design constraints on these factors [179, 253]. Accessibility is also associated with system factors, such as Feedback, because tactile instead of visual feedback may need to be evaluated for users with visual impairments [143]. As this factor, and likely others, can be decomposed into a subset of the 13 factors that this present research identified, it seems important to first establish the set of primitive factors before the inclusion of higher-level factors.

The second factor was multi-user gestural interaction (i.e., the use of gestural interaction in multi-user collaborative or cooperative settings [174, 215, 216]). The factor identification and analysis provided within this article was limited to single user use cases because, as the analysis began to demonstrate, there was yet to be a consensus about all of the factors implicated in the design of single user gestures, let alone a methodology that could evaluate all of the factors at the same time. As the multi-user use case is a more complex version of the single-user scenarios that are commonly evaluated and used as use cases in the literature, it did not seem appropriate to dive into this more complex topic before a solid foundation of single user gestural interaction was developed. While the findings about some factors such as Social Acceptability and Context may seem, at first glance to be the most applicable to these scenarios, there is much need and opportunity to explore how *all* the 13 factors influence these scenarios and also to identify any additional factors that could be unique to multi-user settings (e.g., multi-user cooperation, user roles, territoriality, and so on).

Lastly, one of the most popular factors that emerged was *naturalness*, which has also been referred to as *fluidity*, *feelings*, *satisfaction*, *controllability*, *affect*, and *familiarity*.¹⁴ Naturalness, however, is a complex factor to consider because it is difficult to define and measure objectively. Baudel and Beaudouin-Lafon argued that gestural interaction is *natural* because it builds upon the manipulation and gestural communication skills that humans acquire naturally [37]. Wigdor and

¹⁴See references including [1, 28, 29, 37, 54, 55, 65, 73, 86, 91, 98, 100–103, 108, 110, 115, 116, 119, 134, 157, 169, 170, 175–177, 186, 187, 206, 208, 213, 226, 235, 235, 237, 238, 242, 244, 246, 250, 252, 255, 258, 281, 282, 284, 288, 291, 294, 295, 301, 302, 314, 319, 323, 324, 329, 333, 347, 349, 350, 362, 368, 382, 384].

Wixon, however, proposed that naturalness is not in fact a design factor at all, but rather an experience goal, wherein a system should evoke a feeling of effortlessness and enjoyment that results in users “act[ing] and feel[ing] like a natural” in its use [357]. These “natural” feelings are often compounded by many relevant factors, such as Intuitiveness, Learnability, and Ergonomics. The pinch to zoom gesture, for example, does not mimic or build upon an existing hand motions or metaphors that users are accustomed to, but over repeated presentations and performances of the gesture, has been described as a fluid, satisfying, familiar gesture by users who perform it every-day [357].

Gestural interaction has often been used as one example of a “natural” user interface when compared to the Windows, Icons, Menus, Pointer user interface metaphor [320]. Despite naturalness having been repeatedly referred to in the literature, the definition of naturalness has been inconsistent [37, 357] and was often used alongside descriptions of other factors such as Learnability, Intuitiveness, and Complexity [119, 121, 359, 376]. Given the lack of consensus regarding what constitutes a “natural” movement or feeling and the fact that naturalness is often referred to as an experience goal rather than as a design factor [101, 208, 300], naturalness was not considered to be one of the factors crucial to the design of gestures within the context of this work. We respectfully encourage the community to avoid the use of the terms “natural” or “naturalness” when describing interactive systems, and instead focus on using terms to describe the situational, cognitive, physical, and system factors that govern gestural interaction, such as those that were identified, analyzed, summarized in this work.

6 CONCLUSION

Over the decades, gestural interaction has become an increasingly dominant way to interact with computing devices, such as touchscreens, mobile phones, or watches, or interact with immersive experiences, such as augmented and virtual reality. Despite the widespread adoption of devices with gesture sensing capabilities, the design of gesture vocabularies that suit the various needs and constraints of an application, its context, and its users, is still a challenging task. This work documents the efforts undertaken within the HCI community to design and understand gestural interaction, through the invention of new input modalities and sensing techniques, as well as new ways to measure and evaluate gesture vocabularies.

While a significant amount of work has explored different gesture design methodologies and techniques that can address the problems inherent in gestural interaction, there is a lack of work that has holistically understood, identified, or analyzed the various factors that are essential to consider when designing gestural vocabularies. This work reviewed the literature on gesture design and summarized the three main gesture design methodologies used today, i.e., Expert-led, User-led, and Computationally-based, as well as identified 13 factors, separated into four categories, that were crucial to the design of gestural experiences. These including Situational (i.e., Context, Modality, and Social Acceptability), Cognitive (i.e., Discoverability, Intuitiveness, Learnability, and Transferability), Physical (i.e., Efficiency, Complexity, Ergonomics, and Occlusion), and System (i.e., Feedback and Recognition) factors. Best practices to evaluate and refine gestures were also summarized for each of the identified factors, and a factor-centric analysis of the existing gesture-design methodologies demonstrated that none of the existing methodologies satisfy all the identified factors.

We thus proposed how iterative, user-focused, factor-centric methodology approaches could serve as a practical guide for gesture designers to develop suitable gesture vocabularies in the future. Such methodologies could identify the preconditions that will influence user experiences with the gestures, assign weights to each factor to indicate their relative priority, call for the creation of an initial gesture vocabulary, and encourage designers to continually refine and

evaluate the gesture vocabulary while being mindful of the importance of different factors. Lastly, we concluded with a short discussion on three additional factors that we encountered (i.e., multi-user, accessibility, and naturalness) and discuss how these factors are either compositions of the 13 factors that we identified or are descriptors of experience goals, rather than design factors.

Through these exercises, we believe that this article has provided a foundational understanding of the factors that influence gestural design, how these factors can be evaluated, and the implications that arise when different methodologies do or do not allow for an evaluation of such factors. Although there is much research still to be done within respect to modeling and understanding factors such as Transferability, Feedback, and Intuitiveness, in addition to examining more complex scenarios involving multiple users, this research should help designers, developers, and researchers develop a holistic understanding of the elements and characteristics that govern gestural interaction and should pave the way for much future innovation with respect to the gesture design methodologies, tools, and evaluations that will be possible.

REFERENCES

- [1] Hajar G. H. Abadi, Lim Yan Peng, and Ali Mohammad Hossein Zadeh. 2012. Guessability study on considering cultural values in gesture design for different user interfaces. In *Proceedings of the International Proceedings of Economics Development and Research*, Vol. 37. 1–4.
- [2] Christopher Ackad, Judy Kay, and Martin Tomitsch. 2014. Towards learnable gestures for exploring hierarchical information spaces at a large public display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*, Vol. 49. ACM, New York, NY, 57.
- [3] Anshu Agarwal and Madhu Prabaker. 2009. Building on the usability study: Two explorations on how to better understand an interface. In *Proceedings of the International Conference on Human-Computer Interaction*. ACM, New York, NY, 385–394.
- [4] Anant Agarwal and Manish K. Thakur. 2013. Sign language recognition using microsoft kinect. In *Proceedings of the International Conference on Contemporary Computing*. ACM, New York, NY, 181–185. DOI: doi.org/10.1109/IC3.2013.6612186
- [5] David Ahlström, Khalad Hasan, and Pourang Irani. 2014. Are you comfortable doing that? Acceptance studies of around-device gestures in and for public settings. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'14)*. ACM, New York, NY, 193–202. DOI: <https://doi.org/10.1145/2628363.2628381>
- [6] Roland Aigner, Daniel Wigdor, Hrvoje Benko, Michael Haller, David Lindbauer, Alexandra Ion, Shengdong Zhao, and J. T. K. V. Koh. 2012. *Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI*. Technical Report. Microsoft Research.
- [7] Deepak Akkil, Andrés Lucero, Jari Kangas, Tero Jokela, Marja Salmimaa, and Roope Raisamo. 2016. User expectations of everyday gaze interaction on smartglasses. In *Proceedings of the Nordic Conference on Human-Computer Interaction*. ACM, New York, NY, 1–10.
- [8] Fouad Alallah, Ali Neshati, Nima Sheibani, Yumiko Sakamoto, Andrea Bunt, Pourang Irani, and Khalad Hasan. 2018. Crowdsourcing vs laboratory-style social acceptability studies? Examining the social acceptability of spatial user interactions for head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–7.
- [9] Abdullah X. Ali, Meredith Ringel Morris, and Jacob O. Wobbrock. 2018. Crowdsourcing similarity judgments for agreement analysis in end-user elicitation studies. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 177–188. DOI: <https://doi.org/10.1145/3242587.3242621>
- [10] Abdullah Almaksour, Eric Anquetil, Solen Quiniou, and Mohamed Cheriet. 2010. Personalizable pen-based interface using lifelong learning. In *Proceedings of the International Conference on Frontiers in Handwriting Recognition*. IEEE, New York, NY, 188–193.
- [11] Shahriyar Amini and Yang Li. 2013. Crowdlearner: Rapidly creating mobile recognizers using crowdsourcing. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 163–172.
- [12] Roman Amstutz, Oliver Amft, Brian French, Asim Smailagic, Dan Siewiorek, and Gerhard Troster. 2009. Performance analysis of an HMM-Based gesture recognition using a wristwatch device. In *Proceedings of the International Conference on Computational Science and Engineering*, Vol. 2. IEEE, New York, NY, 303–309.
- [13] Fraser Anderson and Walter F. Bischof. 2013. Learning and performance with gesture guides. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1109–1118.

- [14] Fraser Anderson, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2015. Supporting subtlety with deceptive devices and illusory interactions. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1489–1498.
- [15] Leonardo Angelini, Francesco Carrino, Stefano Carrino, Maurizio Caon, Omar Abou Khaled, Jürgen Baumgartner, Andreas Sonderegger, Denis Lalanne, and Elena Mugellini. 2014. Gesturing on the steering wheel: A user-elicited taxonomy. In *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, 31.
- [16] Roger B. Annenberg and Dale Amon. 1989. A gesture based user interface prototyping system. In *Proceedings of the 2nd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology*. ACM, New York, NY, 127–132.
- [17] Michelle Annett and Walter F. Bischof. 2013. Your left hand can do it too! Investigating intermanual symmetric gesture transfer on touchscreens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1119–1128.
- [18] Michelle Annett, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2011. Medusa: A proximity-aware multi-touch tabletop. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 337–346. DOI : <https://doi.org/10.1145/2047196.2047240>
- [19] Lisa Anthony, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2013. Understanding the consistency of users’ pen and finger stroke gesture articulation. In *Proceedings of Graphics Interface 2013*. Canadian Human–Computer Communications Society, 87–94.
- [20] Lisa Anthony and Jacob O. Wobbrock. 2010. A lightweight multistroke recognizer for user interface prototypes. In *Proceedings of Graphics Interface*. Canadian Human-Computer Communications Society, 245–252.
- [21] Caroline Appert and Shumin Zhai. 2009. Using strokes as command shortcuts: Cognitive benefits and toolkit support. In *Proceedings of the International Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2289–2298.
- [22] Shaikh Shawon Arefin Shimon, Courtney Lutton, Zichun Xu, Sarah Morrison-Smith, Christina Boucher, and Jaime Ruiz. 2016. Exploring non-touchscreen gestures for smartwatches. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 3822–3833.
- [23] Dante Arias-Torres. 2006. The design and evaluation of a pen-based computer interface for novice older users. In *Proceedings of the Mexican International Conference on Computer Science*. ACM, New York, NY, 142–150.
- [24] Ahmed Sabbir Arif and Wolfgang Stuerzlinger. 2014. User adaptation to a faulty unistroke-based text entry technique by switching to an alternative gesture language. In *Proceedings of the Graphics Interface*. Canadian Human–Computer Communications Society, 183–192.
- [25] Rahul Arora, Rubaiat Habib Kazi, Danny M. Kaufman, Wilnot Li, and Karan Singh. 2019. MagicalHands: Mid-Air hand gestures for animating in VR. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 463–477. DOI : <https://doi.org/10.1145/3332165.3347942>
- [26] Daniel Ashbrook and Thad Starner. 2010. MAGIC: A motion gesture design tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2159–2168.
- [27] Ilhan Aslan, Ida Buchwald, Philipp Koytek, and Elisabeth André. 2016. Pen+ mid-air: An exploration of mid-air gestures to complement pen input on tablets. In *Proceedings of the Nordic Conference on Human–Computer Interaction*. ACM, New York, NY, 1–10.
- [28] Cohé Aurélie and Martin Hachet. 2012. Understanding user gestures for manipulating 3D objects from touchscreen inputs. In *Proceedings of the Graphics Interface*. Canadian Human-Computer Communications Society, 157–164.
- [29] Daniel Avrahami, Scott E. Hudson, Thomas P. Moran, and Brian D. Williams. 2001. Guided gesture support in the paper PDA. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 197–198.
- [30] Shiri Azenkot and Shumin Zhai. 2012. Touch behavior with different postures on soft smartphone keyboards. In *Proceedings of the International Conference on Human–Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, 251–260.
- [31] Stefan Bachl, Martin Tomitsch, Christoph Wimmer, and Thomas Grechenig. 2010. Challenges for designing the user experience of multi-touch interfaces. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems Workshop on Engineering Patterns for Multi-Touch Interfaces*. ACM, New York, NY.
- [32] Myroslav Bachynskyi, Gregorio Palmas, Antti Oulasvirta, and Tino Weinkauff. 2015. Informing the design of novel input methods with muscle coactivation clustering. *ACM Transactions on Computer–Human Interaction* 21, 6 (2015), 1–25.
- [33] Gilles Bailly, Thomas Pietrzak, Jonathan Deber, and Daniel J. Wigdor. 2013. Métamorphe: Augmenting hotkey usage with actuated keys. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 563–572.

- [34] Kayne Barclay, Danny Wei, Christof Lutteroth, and Robert Sheehan. 2011. A quantitative quality model for gesture based user interfaces. In *Proceedings of the Australian Computer-Human Interaction Conference*. ACM, New York, NY, 31–39.
- [35] Karim El Batran and Mark D. Dunlop. 2014. Enhancing KLM (keystroke-level model) to fit touch screen mobile devices. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, 283–286.
- [36] Olivier Bau and Wendy E. Mackay. 2008. Octopocus: A dynamic guide for learning gesture-based command sets. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 37–46.
- [37] Thomas Baudel and Michel Beaudouin-Lafon. 1993. Charade: Remote control of objects using free-hand gestures. *Communications of the ACM* 36, 7 (1993), 28–35. DOI: <https://doi.org/10.1145/159544.159562>
- [38] Susanne Bay, Philipp Brauner, Thomas Gossler, and Martina Ziefle. 2014. Intuitive gestures on multi-touch displays for reading radiological images. In *Proceedings of the International Conference on Human-Computer Interaction*. Springer, New York, NY, 22–31.
- [39] Mehmet Aydın Baytaş, Yücel Yemez, and Oğuzhan Özcan. 2014. Hotspotizer: End-user authoring of mid-air gestural interactions. In *Proceedings of the Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM, New York, NY, 677–686.
- [40] Hrvoje Benko and Daniel Wigdor. 2010. Imprecision, inaccuracy, and frustration: The tale of touch input. In *Proceedings of the Tabletops - Horizontal Interactive Displays*. Springer, New York, NY, 978–1.
- [41] Mike Bennett, Kevin McCarthy, Sile O’modhrain, and Barry Smyth. 2011. Simpleflow: Enhancing gestural interaction with gesture prediction, abbreviation and autocompletion. In *Proceedings of the IFIP Conference on Human-Computer Interaction*. Springer, New York, NY, 591–608.
- [42] Xiaojun Bi, Tomer Moscovich, Gonzalo Ramos, Ravin Balakrishnan, and Ken Hinckley. 2008. An exploration of pen rolling for pen-based interaction. In *Proceedings of the 21st annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 191–200.
- [43] Per Bjerre and Andreas K. Pedersen. *A Touch-Level Model for Touch Gesture Efficiency*. http://vbn.aau.dk/ws/files/198144570/MTA_14830_A_Touch_level_Model_for_Touch_Gestures.pdf.
- [44] Sebastian Boring, Marko Jurmu, and Andreas Butz. 2009. Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*. ACM, New York, NY, 161–168.
- [45] Ali Boyali and Naohisa Hashimoto. 2014. Block-sparse representation classification based gesture recognition approach for a robotic wheelchair. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, Vol. 10. IEEE, New York, NY, 1133–1138.
- [46] Ali Boyali, Naohisa Hashimoto, and Osamu Matsumoto. 2015. A signal pattern recognition approach for mobile devices and its application to braking state classification on robotic mobility devices. *Robotics and Autonomous Systems* 72 (2015), 37–47. DOI: <https://doi.org/10.1016/J.Robot.2015.04.008>
- [47] Andrew Bragdon, Rob Deline, Ken Hinckley, and Meredith Ringel Morris. 2011. Code space: Touch + Air gesture hybrid interactions for supporting developer meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 212–221.
- [48] Andrew Bragdon, Eugene Nelson, Yang Li, and Ken Hinckley. 2011. Experimental analysis of touch-screen gesture designs in mobile environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 403–412.
- [49] Andrew Bragdon, Arman Uguray, Daniel Wigdor, Stylianos Anagnostopoulos, Robert Zeleznik, and Rutledge Feman. 2010. Gesture play: Motivating online gesture learning with fun. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 39–48.
- [50] Andrew Bragdon, Robert Zeleznik, Brian Williamson, Timothy Miller, and Joseph J. La Viola, Jr. 2009. GestureBar: Improving the approachability of gesture-based interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2269–2278.
- [51] Peter Brandl, Clifton Forlines, Daniel Wigdor, Michael Haller, and Chia Shen. 2008. Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In *Proceedings of the Working Conference on Advanced Visual Interfaces*. ACM, New York, NY, 154–161.
- [52] Helene Brashear, Thad Starner, Paul Lukowicz, and Holger Junker. 2003. Using multiple sensors for mobile sign language recognition. In *Proceedings of the IEEE International Symposium on Wearable Computers*. IEEE, New York, NY, 45–52.
- [53] John Brooke. 1996. Sus: A quick and dirty usability. *Usability Evaluation in Industry* 189, 194 (1996), 4–7.
- [54] Sarah Buchanan, Bourke Floyd, Will Holderness, and Joseph J. La Viola Jr. 2013. Towards user-defined multi-touch gestures for 3D objects. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 231–240.

- [55] Gary Burnett, Elizabeth Crundall, David Large, Glyn Lawson, and Lee Skrypchuk. 2013. A study of unidirectional swipe gestures on in-vehicle touch screens. In *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, 22–29.
- [56] Bill Buxton. 1995. Chunking and phrasing and the design of human–computer dialogues. *Readings in Human–Computer Interaction* (1995), 494–499.
- [57] Marcio C. Cabral, Carlos H. Morimoto, and Marcelo K. Zuffo. 2005. On the usability of gesture interfaces in virtual reality environments. In *Proceedings of the 2005 Latin American conference on Human–Computer Interaction*. 100–108.
- [58] Francesco Cafaro, Leilah Lyons, Raymond Kang, Josh Radinsky, Jessica Roberts, and Kristen Vogt. 2014. Framed guessability: Using embodied allegories to increase user agreement on gesture languages. In *Proceedings of the International Conference on Tangible Embedded and Embodied Interaction*. ACM, New York, NY, 197–204.
- [59] Xiang Cao and Shumin Zhai. 2007. Modeling human performance of pen stroke gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1495–1504.
- [60] Jessica R. Cauchard, L. E. Jane, Kevin Y. Zhai, and James A. Landay. 2015. Drone and me: An exploration into natural human-drone interaction. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, 361–365.
- [61] Hsien-Tsung Chang, Tsai-Hsuan Tsai, Ya-Ching Chang, and Yi-Min Chang. 2014. Touch panel usability of elderly and children. *Computer Human Behavior* 37 (2014), 258–269. DOI : <https://doi.org/10.1016/J.Chb.2014.04.050>
- [62] Li-Chieh Chen, Po-Ying Chu, and Yun-Maw Cheng. 2016. Exploring the ergonomic issues of user-defined mid-air gestures for interactive product exhibition. In *Proceedings of the International Conference on Distributed Ambient, and Pervasive Interactions*. 180–190.
- [63] Guojun Cheng and Jibin Yin. 2013. The study of a novel gesture technique based on “Pen+ touch” in multi-touch interfaces. In *Proceedings of the 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer*. 1130–1134.
- [64] Sung-Jung Cho, Jong Koo Oh, Won-Chul Bang, Wook Chang, Eunseok Choi, Yang Jing, Joonkee Cho, and Dong Yoon Kim. 2004. Magic wand: A hand-drawn gesture input device in 3-D space with inertial sensors. In *Proceedings of the 9th International Workshop on Frontiers in Handwriting Recognition*. 106–111. DOI : <http://dx.doi.org/10.1109/TWFHR.2004.66>
- [65] Eunjung Choi, Sunghyuk Kwon, Donghun Lee, Hogin Lee, and Min K. Chung. 2014. Towards successful user interaction with systems: Focusing on user-derived gestures for smart home systems. In *Proceedings of the Applied Ergonomics Conference*. 1196–1207. DOI : <http://dx.doi.org/j.apergo.2014.02.010>
- [66] Andy Cockburn, David Ahlström, and Carl Gutwin. 2012. Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human–Computer Studies* 70, 3 (2012), 218–233.
- [67] Andy Cockburn, Carl Gutwin, Joey Scarr, and Sylvain Malacria. 2014. Supporting novice to expert transitions in user interfaces. *ACM Computing Surveys* 47, 2 (2014), 31:1–31:36. DOI : <https://doi.org/10.1145/2659796>
- [68] Gabe Cohn, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Humantenna: Using the body as an antenna for real-time whole-body interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1901–1910.
- [69] Sabrina Connell, Pei-Yi Kuo, Liu Liu, and Anne Marie Piper. 2013. A wizard-of-oz elicitation study examining child-defined gestures with a whole-body interface. In *Proceedings of the International Conference on Interaction Design and Children*. ACM, New York, NY, 277–280.
- [70] Nancy J. Cooke. 1994. Varieties of knowledge elicitation techniques. *International Journal of Human–Computer Studies* 41, 6 (1994), 801–849. DOI : <https://doi.org/10.1006/ijhc.1994.1083>
- [71] Gabriele Costante, Lorenzo Porzi, Oswald Lanz, Paolo Valigi, and Elisa Ricci. 2014. Personalizing a smartwatch-based gesture interface with transfer learning. In *Proceedings of the European Signal Processing Conference*. IEEE, New York, NY, 2530–2534.
- [72] Joelle Coutaz, Laurence Nigay, Daniel Salber, Ann Blandford, Jon May, and Richard M. Young. 1995. Four easy pieces for assessing the usability of multimodal interaction: The CARE properties. *Human–Computer Interaction* (1995), 115–120.
- [73] Florian Daiber, Johannes Schöning, and Antonio Krüger. 2011. Towards a framework for whole body interaction with geospatial data. *Whole Body Interaction* 10, 1007 (2011), 197–207. DOI : https://doi.org/978-0-3_16
- [74] Chi Tai Dang and Elisabeth André. 2011. Usage and recognition of finger orientation for multi-touch tabletop interaction. In *Proceedings of the IFIP Conference on Human–Computer Interaction*. Springer, 409–426.
- [75] William Delamare, Céline Coutrix, and Laurence Nigay. 2015. Designing guiding systems for gesture-based interaction. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. ACM, New York, NY, 44–53.

- [76] William Delamare, Thomas Janssoone, Céline Coutrix, and Laurence Nigay. 2016. Designing 3D gesture guidance: Visual feedback and feedforward design options. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. 152–159.
- [77] William Delamare, Chaklam Silpasuwanchai, Sayan Sarcar, Toshiaki Shiraki, and Xiangshi Ren. 2020. On gesture combination: An exploration of a solution to augment gesture interaction. *Proceedings of the ACM on Human-Computer Interaction* 4, EICS (2020), 1–21.
- [78] Stefano Dessi and Lucio Davide Spano. 2020. DG3: Exploiting gesture declarative models for sample generation and online recognition. *Proceedings of the ACM on Human-Computer Interaction* 4, EICS (2020), 1–21.
- [79] Anind K. Dey. 2001. Understanding and using context. *Personal Ubiquitous Computing* 5, 1 (2001), 4–7. DOI : <https://doi.org/10.1007/s007790170019>
- [80] Travis Deyle, Szabolcs Palinko, Erika Shehan Poole, and Thad Starner. 2007. Hambone: A bio-acoustic gesture interface. In *Proceedings of the IEEE International Symposium on Wearable Computers*. IEEE, New York, NY, 3–10.
- [81] D. Diaper and N. Stanton. 2003. *The handbook of task analysis for human-computer interaction*. CRC Press.
- [82] Nem Khan Dim, Chaklam Silpasuwanchai, Sayan Sarcar, and Xiangshi Ren. 2016. Designing Mid-Air TV gestures for blind people using user-and choice-based elicitation approaches. In *Proceedings of the ACM Conference on Designing Interactive Systems*. ACM, New York, NY, 204–214.
- [83] Alan Dix, Janet Finlay, Gregory Abowd, and Russell Beale. 2004. “Evaluation techniques.” *Human-computer interaction* (2004), 318–364.
- [84] Haiwei Dong, Ali Danesh, Nadia Figueroa, and Abdulmotaleb El Saddik. 2015. An elicitation study on gesture preferences and memorability toward a practical hand-gesture vocabulary for smart televisions. *IEEE Access* 3 (2015), 543–555.
- [85] Ze Dong, Thammathip Piumsomboon, Jingjing Zhang, Adrian Clark, Huidong Bai, and Rob Lindeman. 2020. A comparison of surface and motion user-defined gestures for mobile augmented reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New York, NY, 1–8.
- [86] Tanja Döring, Dagmar Kern, Paul Marshall, Max Pfeiffer, Johannes Schöning, Volker Gruhn, and Albrecht Schmidt. 2011. Gestural interaction on the steering wheel: Reducing the visual demand. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 483–492.
- [87] Abdallah El Ali, Johan Kildal, and Vuokko Lantz. 2012. Fishing on a Z?: Investigating the effects of error on mimetic and alphabet device-based gesture interaction. In *Proceedings of the ACM International Conference on Multimodal Interaction*. ACM, New York, NY, 93–100.
- [88] Jerry Fails and Dan Olsen. 2003. A design tool for camera-based interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 449–456.
- [89] Farzin Farhadi-Niaki, S. Ali Etamad, and Ali Arya. 2013. Design and usability analysis of gesture-based control for common desktop tasks. In *Proceedings of the International Conference on Human-Computer Interaction*. Springer, 215–224.
- [90] Wim Fikkert, Paul van der Vet, and Anton Nijholt. 2010. User-evaluated gestures for touchless interactions from a distance. In *Proceedings of the IEEE International Symposium on Multimedia*. IEEE, New York, NY, 153–160.
- [91] Wim Fikkert, Paul Van Der Vet, Gerrit van der Veer, and Anton Nijholt. 2009. Gestures for large display control. *International Gesture Workshop* 10, 1007 (2009), 245–256. DOI : https://doi.org/10.1007/978-3-9_22
- [92] Clifton Forlines, Daniel Wigdor, Chia Shen, and Ravin Balakrishnan. 2007. Direct-touch vs. mouse input for tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 647–656.
- [93] Dustin Freeman, Hrvoje Benko, Meredith Ringel Morris, and Daniel Wigdor. 2009. Shadowguides: Visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 165–172.
- [94] Dustin Freeman, Ramadevi Vennelakanti, and Sriganesh Madhvanath. 2012. Freehand pose-based gestural interaction: Studies and implications for interface design. In *Proceedings of the International Conference on Intelligent Human Computer Interaction*, 1–6. doi.org/10.1109/IHCI.2012.6481816
- [95] Mathias Frisch, Jens Heydekorn, and Raimund Dachselt. 2009. Investigating multi-touch and pen gestures for diagram editing on interactive surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 149–156.
- [96] Mathias Frisch, Jens Heydekorn, and Raimund Dachselt. 2010. Diagram editing on interactive displays using multi-touch and pen gestures. In *Proceedings of the International Conference on Diagrammatic Representation and Inference*. 182–196. DOI : https://doi.org/10.1007/978-3-642-14600-8_18
- [97] Qin Gao and Qiqi Sun. 2015. Examining the usability of touch screen gestures for older and younger adults. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 5 (2015), 835–863. DOI : <https://doi.org/10.1177/0018720815581293>

- [98] Franca Garzotto and Matteo Valoriani. 2013. Touchless gestural interaction with small displays: A case study. In *Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI*.
- [99] Vito Gentile, Alessio Malizia, Salvatore Sorce, and Antonio Gentile. 2015. Designing touchless gestural interactions for public displays in-the-wild. In *Proceedings of the International Conference on Human-Computer Interaction*. 24–34.
- [100] Brian Gleeson, Karon MacLean, Amir Haddadi, Elizabeth Croft, and Javier Alcazar. 2013. Gestures for industry: Intuitive human-robot communication from human observation. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. IEEE, New York, NY, 349–356.
- [101] Sureshini A. Grandhi, Gina Joue, and Irene Mittelberg. 2011. Understanding naturalness and intuitiveness in gesture production: Insights for touchless gestural interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 821–824.
- [102] Sureshini A. Grandhi, Gina Joue, and Irene Mittelberg. 2012. To move or to remove?: A human-centric approach to understanding gesture interpretation. In *Proceedings of the Designing Interactive Systems Conference*. ACM, New York, NY, 582–591.
- [103] Sureshini A. Grandhi, Chat Wacharamanatham, Gina Joue, Jan O. Borchers, and Irene Mittelberg. 2013. How we gesture towards machines: An exploratory study of user perceptions of gestural interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New York, NY, 1209–1214.
- [104] Daniela Grijincu, Miguel A. Nacenta, and Per Ola Kristensson. 2014. User-Defined interface gestures: Dataset and analysis. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 25–34.
- [105] Robin Guérit, Alessandro Cierro, Jean Vanderdonckt, and Jorge Luis Pérez-Medina. 2019. Gesture elicitation and usability testing for an armband interacting with netflix and spotify. In *Proceedings of the International Conference on Information Technology and Systems*. Springer, 625–637.
- [106] Anind K. Dey and Raffay Hamid, Chris Beckmann, Ian Li, and Daniel Hsu. 2004. a CAPpella: Programming by demonstration of context-aware applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 33–40.
- [107] Jeffrey T. Hansberger, Chao Peng, Shannon L. Mathis, Vaidyanath Areyur Shanthakumar, Sarah C. Meacham, Lizhou Cao, and Victoria R. Blakely. 2017. Dispelling the gorilla arm syndrome: The viability of prolonged gesture interactions. In *Proceedings of the International Conference on Virtual, Augmented and Mixed Reality*. Springer, 505–520.
- [108] Philip A. Harling and Alistair D. N. Edwards. 1996. Hand tension as a gesture segmentation cue. In *Proceedings of the Gesture Workshop on Progress in Gestural Interaction*. 75–88.
- [109] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the body as an input surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 453–462.
- [110] Chris Harrison, Robert Xiao, Julia Schwarz, and Scott E. Hudson. 2014. TouchTools: Leveraging familiarity and skill with physical tools to augment touch interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2913–2916.
- [111] Björn Hartmann, Leith Abdulla, Manas Mittal, and Scott R. Klemmer. 2007. Authoring sensor-based interactions by demonstration with direct manipulation and pattern recognition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 145–154.
- [112] Seongkook Heo, Michelle Annett, Benjamin Lafreniere, Tovi Grossman, and George Fitzmaurice. 2017. No need to stop what you're doing: Exploring no-handed smartwatch interaction. In *Proceedings of the Graphics Interface Conference*. Canadian Human-Computer Communications Society, 107–114.
- [113] Marc Heseniuss and Volker Gruhn. 2016. Introducing gesturecards: A prototyping gesture notation. In *Proceedings of the Nordic Conference on Human-Computer Interaction*. ACM, New York, NY.
- [114] Tobias Hesselmann and Susanne Boll. 2010. SCiVA: A design process for applications on interactive surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 265–266.
- [115] Jens Heydekorn, Mathias Frisch, and Raimund Dachselt. 2011. Evaluating a user-elicited gesture language for interactive displays. In *Proceedings of the Mensch and Computer*. 191–200.
- [116] Otmar Hilliges, Andreas Butz, Shahram Izadi, and Andrew D. Wilson. 2010. Interaction on the tabletop: Bringing the physical to the digital. In *Tabletops - Horizontal Interactive Displays*, C. Müller-Tomfelde (Ed.). Human-Computer Interaction Series. Springer, London. <https://doi.org/10.1007/978-1-84996-113>
- [117] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1063–1072.
- [118] Ken Hinckley, Patrick Baudisch, Gonzalo Ramos, and Francois Guimbretiere. 2005. Design and analysis of delimiters for selection-action pen gesture phrases in scriboli. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 451–460.

- [119] Ken Hinckley, Michel Pahud, Hrvoje Benko, Pourang Irani, François Guimbrètière, Marcel Gavrilu, Xiang 'Anthony' Chen, Fabrice Matulic, William Buxton, and Andrew Wilson. 2014. Sensing techniques for tablet+stylus interaction. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 605–614.
- [120] Ken Hinckley and Hyunyoung Song. 2011. Sensor synaesthesia: Touch in motion and motion in touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 801–810.
- [121] Ken Hinckley, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko, and Bill Buxton. 2010. Pen + Touch = New Tools. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 27–36. DOI : <https://doi.org/10.1145/1866029.1866036>
- [122] Uta Hinrichs and Sheelagh Carpendale. 2011. Gestures in the wild: Studying multi-touch gesture sequences on interactive tabletop exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 3023–3032.
- [123] Thorsten Hochreuter. 2014. LemonGrasp: A tool for touch-interaction prototyping. In *Proceedings of the International Conference on Tangible, Embedded and Embodied Interaction*. ACM, New York, NY, 317–320.
- [124] Michael Hoffman, Paul Varcholik, and Joseph J. La Viola. 2010. Breaking the status quo: Improving 3D gesture recognition with spatially convenient input devices. In *Proceedings of the IEEE Virtual Reality Conference*. IEEE, New York, NY, 59–66. DOI : <https://doi.org/10.1109/VR.2010.5444813>
- [125] Eve Hoggan, Miguel Nacenta, Per Ola Kristensson, John Williamson, Antti Oulasvirta, and Anu Lehtiö. 2013. Multi-touch pinch gestures: Performance and ergonomics. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 219–222.
- [126] Steven Hooper and Eric Berkman. 2011. *Designing Mobile Interfaces*. O'Reilly Media Inc.
- [127] Chi-Yuan Hu, Heng-Yi Lin, and Li-Chieh Chen. 2016. The effects of screen size on rotating 3D contents using compound gestures on a mobile device. *Displays* 41 (2016), 25–32.
- [128] Caroline Hummels and Pieter Jan Stappers. 1998. Meaningful gestures for human computer interaction: Beyond hand postures. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, New York, NY, 591–596. DOI : doi.org/10.1109/AFGR.1998.671012
- [129] Poika Isokoski. 2001. Model for unistroke writing time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 357–364.
- [130] Robert J. K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: A framework for post-WIMP interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 201–210.
- [131] Ivar Jacobson. 1992. *Object-oriented software engineering: A use case driven approach*. Addison-Wesley, Reading, MA.
- [132] Sujin Jang, Niklas Elmquist, and Karthik Ramani. 2014. Gestureanalyzer: Visual analytics for pattern analysis of mid-air hand gestures. In *Proceedings of the ACM Symposium on Spatial User Interaction*. ACM, New York, NY, 30–39.
- [133] Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. 2017. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 3328–3339.
- [134] William Judd, Pourang Irani, Sriram Subramanian, Jason Alexander, and Teng Han. 2012. Putting your best foot forward: Investigating real-world mappings for foot-based gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1229–1238.
- [135] 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*. IEEE, New York, NY, 7–10.
- [136] Hans-Christian Jetter, Svenja Leifert, Jens Gerken, Sören Schubert, and Harald Reiterer. 2012. Does (multi-) touch aid users' spatial memory and navigation in 'panning' and 'zooming and panning' UIs?. In *Proceedings of the International Conference on Advanced Visual Interfaces*. ACM, New York, NY, 83–90.
- [137] Youngkyu Jin, Seungeok Choi, Amy Chung, Insik Myung, Jayeon Lee, Minchul Kim, and Jookyung Woo. 2004. GIA: Design of a gesture-based interaction photo album. *Personal Ubiquitous Computing* 8, 3–4 (2004), 227–233. DOI : <https://doi.org/10.1007/S00779-004-0282-Y>
- [138] Markus Jokisch, Thomas Bartoschek, and Angela Schwering. 2011. Usability testing of the interaction of novices with a multi-touch table in semi public space. In *Proceedings of the International Conference on Human-Computer Interaction: Interaction Techniques and Environments - Volume Part II*. 71–80. DOI : https://doi.org/10.1007/978-3-642-21605-3_8
- [139] Alistair Jones. 2013. *Co-located collaboration in interactive spaces for preliminary design*. PhD Dissertation. Universite de Technologie de Compiegne.
- [140] Ricardo Jota, Pedro Lopes, Daniel Wigdor, and Joaquim Jorge. 2014. Let's kick it: How to stop wasting the bottom third of your large screen display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1411–1414. DOI : <https://doi.org/10.1145/2556288.2557316>

- [141] Sanna Kallio, Juha Kela, and Jani Mantyjarvi. 2003. Online gesture recognition system for mobile interaction. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics 3*. IEEE, New York, NY, 2070–2076.
- [142] Shaun K. Kane, Meredith Ringel Morris, Annuska Z. Perkins, Daniel Wigdor, Richard E. Ladner, and Jacob O. Wobbrock. 2011. Access overlays: Improving non-visual access to large touch screens for blind users. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 273–282.
- [143] Shaun K. Kane, Meredith Ringel Morris, and Jacob O. Wobbrock. 2013. Touchplates: Low-cost tactile overlays for visually impaired touch screen users. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, 1–8.
- [144] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable gestures for blind people: Understanding preference and performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 413–422.
- [145] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable gestures for blind people: Understanding preference and performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 413–422.
- [146] Maria Karam and M. C. Schraefel. 2005. A Taxonomy of Gestures in Human Computer Interactions. Project Report. <https://eprints.soton.ac.uk/261149/>
- [147] Amy K. Karlson, Benjamin B. Bederson, and John Sangiovanni. 2005. AppLens and launchtile: Two designs for one-handed thumb use on small devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 201–210.
- [148] Jun Kato, Sean McDermid, and Xiang Cao. 2012. DeJaVu: Integrated support for developing interactive camera-based programs. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 189–196.
- [149] Keiko Katsuragawa, Ankit Kamal, and Edward Lank. 2017. Effect of motion-gesture recognizer error pattern on user workload and behavior. In *Proceedings of the International Conference on Intelligent User Interfaces*. ACM, New York, NY, 439–449.
- [150] Rafiqul Zaman Khan and Noor Adnan Ibraheem. 2012. Hand gesture recognition: A literature review. *International Journal of Artificial Intelligence and Applications* 3, 161 (2012), 4.
- [151] Jungsoo Kim, Jiasheng He, Kent Lyons, and Thad Starner. 2007. The gesture watch: A wireless contact-free gesture based wrist interface. In *Proceedings of the IEEE International Symposium on Wearable Computers*. IEEE, New York, NY, 1–8.
- [152] Ju-Whan Kim and Tek-Jin Nam. 2013. EventHurdle: Supporting designers’ exploratory interaction prototyping with gesture-based sensors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 267–276.
- [153] J. W. Kim, T. J. Nam, and T. Park. 2017. CompositeGesture: Creating custom gesture interfaces with multiple mobile or wearable devices. *International Journal on Interactive Design and Manufacturing* 11, 1 (2017), 77–82.
- [154] Sang-Hwan Kim, Joonho Ok, Hyun Joo Kang, Min-Chul Kim, and Mijeong Kim. 2004. An interaction and product design of gesture based TV remote control. In *Proceedings of the ACM Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New York, NY, 1548–1548.
- [155] Tony Soo Kim, Sungjae Hwang, and Junghyeon Gim. 2015. SNAP: Sensor aid prototyping tool for designers. In *Proceedings of the ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, 2217–2222.
- [156] Woojoo Kim and Shuping Xiong. 2021. User-Defined walking-in-place gestures for VR locomotion. *International Journal of Human-Computer Studies* 153 (2021), 102648.
- [157] Felix Kistler, Elisabeth André, Samuel Mascarenhas, André Silva, Ana Paiva, Nick Degens, Gert Jan Hofstede, Eva Krumhuber, Arvid Kappas, and Ruth Aylett. 2013. Traveller: An interactive cultural training system controlled by user-defined body gestures. In *Proceedings of the Conference on Human-Computer Interaction*. ACM, New York, NY, 697–704.
- [158] Scott R. Klemmer, Jack Li, James Lin, and James A. Landay. 2004. Papier-mache: Toolkit support for tangible input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 399–406.
- [159] Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. 2020. Social acceptability in hci: A survey of methods, measures, and design strategies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–19.
- [160] Daniel Kohlsdorf, Thad Starner, and Daniel Ashbrook. 2011. MAGIC 2.0: A web tool for false positive prediction and prevention for gesture recognition systems. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition and Workshops*. 1–6.
- [161] Jens Kolb, Benjamin Rudner, and Manfred Reichert. 2013. Gesture-Based process modeling using multi-touch devices. *International Journal on Information System Modelling and Design* 4, 4 (2013), 48–69. DOI: <https://doi.org/10.4018/Ijismd.2013100103>

- [162] Anne Köpsel, Päivi Majaranta, Poika Isokoski, and Anke Huckauf. 2016. Effects of auditory, haptic and visual feedback on performing gestures by gaze or by hand. *Behaviour and Information Technology* 35, 12 (2016), 1044–1062.
- [163] Sven Kratz and Rafael Ballagas. 2007. Gesture recognition using motion estimation on mobile phones. In *Proceedings of the International Workshop on Pervasive Mobile Interaction Devices*. ACM, New York, NY.
- [164] Sven Kratz and Rafael Ballagas. 2009. Unravelling seams: Improving mobile gesture recognition with visual feedback techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 937–940.
- [165] Sven Kratz and Michael Rohs. 2011. Protractor3D: A closed-form solution to rotation-invariant 3D gestures. In *Proceedings of the International Conference on Intelligent User Interfaces*. ACM, New York, NY, 371–374.
- [166] Per Ola Kristensson and L. C. Denby. 2011. Continuous recognition and visualization of pen strokes and touch-screen gestures. In *Proceedings of the Eurographics Symposium on Sketch-Based Interfaces and Modeling*. Eurographics Association, 95–102.
- [167] Per Ola Kristensson, Thomas Nicholson, and Aaron Quigley. 2012. Continuous recognition of one-handed and two-handed gestures using 3D full-body motion tracking sensors. In *Proceedings of the ACM International Conference on Intelligent User Interfaces*. ACM, New York, NY, 89–92.
- [168] Per-Ola Kristensson and Shumin Zhai. 2004. SHARK2: A large vocabulary shorthand writing system for pen-based computers. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 43–52. DOI: <https://doi.org/10.1145/1029632.1029640>
- [169] Christine Kuhnel, Tilo Westermann, Fabian Hemmert, Sven Kratz, Alexander Muller, and Sebastian M. Oller. 2011. I'm Home: Defining and evaluating a gesture language for smart-home control. *International Journal of Human-Computer Studies* 69, 11 (2011), 693–704. DOI: <https://doi.org/10.1016/j.ijhcs.2011.04.005>
- [170] 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*. ACM, New York, NY, 93–96.
- [171] Gord Kurtenbach and Eric Hulteen. 1990. Gestures in human-computer communications. *The Art of Human Computer Interface Design* (1990), 309–317.
- [172] James A. Landay and Brad A. Myers. 1993. Extending an existing user interface toolkit to support gesture recognition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Companion*. ACM, New York, NY, 91–92.
- [173] Songyang Lao, Xiangang Heng, Guohua Zhang, Yunxiang Ling, and Peng Wang. 2009. A gestural interaction design model for multi-touch displays. In *Proceedings of the British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*. British Computer Society, 440–446.
- [174] Jean-Yves Lionel Lawson, Jean Vanderdonck, and Radu-Daniel Vatavu. 2018. Mass-computer interaction for thousands of users and beyond. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New York, NY, 1–6.
- [175] Bongshin Lee, Petra Isenberg, Nathalie Henry Riche, and Sheelagh Carpendale. 2012. Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2689–2698. DOI: <https://doi.org/10.1109/TVCG.2012.204>
- [176] Sang-Su Lee, Jeonghun Chae, Hyunjeong Kim, Younkyung Lim, and Kunpyo Lee. 2013. Towards more natural digital content manipulation via user freehand gestural interaction in a living room. In *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, 617–626.
- [177] Sang-Su Lee, Sohyun Kim, Bopil Jin, Eunji Choi, Boa Kim, Xu Jia, Daeop Kim, and Kun-Pyo Lee. 2010. How users manipulate deformable displays as input devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1647–1656.
- [178] Sang-Su Lee, Youn-Kyung Lim, and Kun-Pyo Lee. 2012. Exploring the effects of size on deformable user interfaces. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services Companion*. ACM, New York, NY, 89–94.
- [179] Roxanne Leitao and Paula Alexandra Silva. 2013. A study of novice older adults and gestural interaction on smartphones. In *Proceedings of the Workshop on Mobile Accessibility*. 1–4.
- [180] L. A. Leiva, D. Martín-Albo, and R. Plamondon. 2016. Gestures à go go: Authoring synthetic human-like stroke gestures using the kinematic theory of rapid movements. *ACM Transactions on Intelligent Systems and Technology* 7, 2 (2016), 15.
- [181] Luis A. Leiva, Daniel Martín-Albo, Réjean Plamondon, and Radu-Daniel Vatavu. 2018. KeyTime: Super-accurate prediction of stroke gesture production times. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1–12.
- [182] Luis A. Leiva, Daniel Martín-Albo, and Radu-Daniel Vatavu. 2017. Synthesizing stroke gestures across user populations: A case for users with visual impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 4182–4193.

- [183] Luis A. Leiva, Daniel Martín-Albo, and Radu-Daniel Vatavu. 2018. GATO: Predicting human performance with multistroke and multitouch gesture input. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–11.
- [184] Luis A. Leiva, Daniel Martín-Albo, Radu-Daniel Vatavu, and Réjean Plamondon. 2021. Stroke gesture synthesis in human-computer interaction. In *Proceedings of the Lognormality Principle and its Applications in E-Security, E-Learning and E-Health*. World Scientific, 211–235.
- [185] Luis A. Leiva, Radu-Daniel Vatavu, Daniel Martín-Albo, and Réjean Plamondon. 2020. Omnis Prædictio: Estimating the full spectrum of human performance with stroke gestures. *International Journal of Human-Computer Studies* 142 (2020), 102466.
- [186] Hai-Ning Liang, Cary Williams, Myron Semegen, Wolfgang Stuerzlinger, and Pourang Irani. 2012. User-defined surface+motion gestures for 3D manipulation of objects at a distance through a mobile device. In *Proceedings of the Asia Pacific Conference on Computer Human Interaction*. 299–308.
- [187] Hai-Ning Liang, Cary Williams, Myron Semegen, Wolfgang Stuerzlinger, and Pourang Irani. 2013. An investigation of suitable interactions for 3D manipulation of distant objects through a mobile device. *International Journal of Innovative Computing, Information, and Control* 9, 12 (2013), 4737–4752.
- [188] Ji Hyoun Lim, Chunik Jo, and Dae-Hoon Kim. 2012. Analysis on user variability in gesture interaction. In *Proceedings of the International Conference on Hybrid Information Technology*. 295–302.
- [189] Jaye Nias, Lisa Anthony, Quincy Brown, and Berthel Tate. 2013. Examining the need for visual feedback during gesture interaction on mobile touchscreen devices for kids. In *Proceedings of the International Conference on Interaction Design and Children*. ACM, New York, NY, 157–164.
- [190] Frieder Loch. 2012. *Hierarchical gestures: Gestural shortcuts for touchscreen devices*. Master's thesis. University of Twente.
- [191] Andreas Löcken, Tobias Hesselmann, Martin Pielot, Niels Henze, and Susanne Boll. 2012. User-centered process for the definition of free-hand gestures applied to controlling music playback. *Multimedia Systems* 18, 1 (2012), 15–31.
- [192] Sebastian Loehmann, Martin Knobel, Melanie Lamara, and Andreas Butz. 2013. Culturally independent gestures for in-car interactions. In *Proceedings of the Conference on Human-Computer Interaction*. ACM, New York, NY, 538–545.
- [193] A. Chris Long. 1998. Improving gestures and interaction techniques for pen-based user interfaces. In *Proceedings of the CHI 98 Conference Summary on Human Factors in Computing Systems*. ACM, New York, NY, 58–59.
- [194] A. Chris Long. 2001. *Quill: A gesture design tool for pen-based user interfaces*. Ph.D. Dissertation. University of California, Berkeley.
- [195] A. Chris Long, James A. Landay, and Lawrence A. Rowe. 1999. Implications for a gesture design tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 40–47.
- [196] A. Chris Long, James A. Landay, and Lawrence A. Rowe. 2001. Those look similar! Issues in automating gesture design advice. In *Proceedings of the Workshop on Perceptive User Interfaces*. ACM, New York, NY, 1–5.
- [197] A. Chris Long, James A. Landay, Lawrence A. Rowe, and Joseph Michiels. 2000. Visual similarity of pen gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 360–367.
- [198] Hao Lü and Yang Li. 2012. Gesture coder: A tool for programming multi-touch gestures by demonstration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2875–2884.
- [199] Hao Lü and Yang Li. 2013. Gesture studio: Authoring multi-touch interactions through demonstration and declaration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 257–266.
- [200] Naveen Madapana, Glebys Gonzalez, and Juan Wachs. 2020. Gesture agreement assessment using description vectors. In *Proceedings of the 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, 40–44.
- [201] Shahzad Malik and Joe Laszlo. 2004. Visual touchpad: A two-handed gestural input device. In *Proceedings of the International Conference on Multimodal Interfaces*. 289–296.
- [202] Cristina Manresa-Yee, Esperança Amengual, and Pere Ponsa Asensio. 2013. Usability of vision-based interfaces. In *Proceedings of the Actas del XIV Congreso Internacional Interacción Persona Ordenador*. 113–118.
- [203] Nicolai Marquardt, Johannes Kiemer, and Saul Greenberg. 2010. What caused that touch?: Expressive interaction with a surface through fiduciary-tagged gloves. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 139–142.
- [204] Francisco J. Martínez-Ruiz, Sebastian F. Rauh, and Gerrit Meixner. 2019. Understanding peripheral audiences: From subtle to full body gestures. In *Proceedings of the International Conference on Human Interaction and Emerging Technologies*. Springer, 489–495.
- [205] Fabrice Matulic and Daniel Vogel. 2018. Multiray: Multi-finger raycasting for large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–13.
- [206] Dan Mauney, Jonathan Howarth, Andrew Wirtanen, and Miranda Capra. 2010. Cultural similarities and differences in user-defined gestures for touchscreen user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New York, NY, 4015–4020.

- [207] Winograd Maynes-Aminzade. 2007. Eyepatch: Prototyping camera-based interaction through examples. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY.
- [208] Mark Micire, Munjal Desai, Amanda Courtemanche, Katherine M. Tsui, and Holly A. Yanco. 2009. Analysis of natural gestures for controlling robot teams on multi-touch tabletop surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 41–48.
- [209] Margaret R. Minsky. 1984. Manipulating simulated objects with real-world gestures using a force and position sensitive screen. In *Proceedings of the Conference on Computer Graphics and Interactive Techniques*. ACM, New York, NY, 195–203.
- [210] Gourav Modanwal and Kishor Sarawadekar. 2017. A new dactylogy and interactive system development for blind-computer interaction. *IEEE Transactions on Human-Machine Systems* 48, 2 (2017), 207–212.
- [211] Calkin S. Montero, Jason Alexander, Mark T. Marshall, and Sriram Subramanian. 2010. Would you do that?: Understanding social acceptance of gestural interfaces. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 275–278.
- [212] Meredith Ringel Morris. 2012. Web on the wall: Insights from a multimodal interaction elicitation study. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*. 95–104.
- [213] Meredith Ringel Morris. 2012. Web on the wall: Insights from a multimodal interaction elicitation study. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 95–104.
- [214] 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.
- [215] Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, and Terry Winograd. 2006. Cooperative gestures: Multi-User gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1201–1210.
- [216] Meredith Ringel Morris, Jarrod Lombardo, and Daniel Wigdor. 2010. WeSearch: Supporting collaborative search and sensemaking on a tabletop display. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*. ACM, New York, NY, 401–410.
- [217] Mehrnaz Mostafapour and Mark Hancock. 2014. Exploring narrative gestures on digital surfaces. In *Proceedings of the 9th ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 5–14.
- [218] Lilian Genaro Motti, Nadine Vigouroux, and Philippe Gorce. 2013. Interaction techniques for older adults using touchscreen devices: A literature review. In *Proceedings of the Conference on l'Interaction Homme-Machine*. 125–135.
- [219] Michael Moyle and Andy Cockburn. 2002. Analysing mouse and pen flick gestures. In *Proceedings of the SIGCHI-NZ Symposium on Computer-Human Interaction*. 19–24.
- [220] Florian Müller, Joshua McManus, Sebastian Günther, Martin Schmitz, Max Mühlhäuser, and Markus Funk. 2019. Mind the tap: Assessing foot-taps for interacting with head-mounted displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'19)*. ACM, New York, NY, 1–13. DOI: <https://doi.org/10.1145/3290605.3300707>
- [221] Brita Munsinger and John Quarles. 2019. Augmented reality for children in a confirmation task: Time, fatigue, and usability. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. ACM, New York, NY, 1–5.
- [222] G. R. S. Murthy and R. S. Jadon. 2009. A review of vision based hand gestures recognition. *International Journal of Information Technology and Knowledge Management* 2, 2 (2009), 405–410.
- [223] Sundar Murugappan, Niklas Elmquist, and Karthik Ramani. 2012. Extended multitouch: Recovering touch posture and differentiating users using a depth camera. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 487–496.
- [224] Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture languages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1099–1108.
- [225] Rajalakshmi Nandakumar, Bryce Kellogg, and Shyamnath Gollakota. 2014. Wi-fi gesture recognition on existing devices. arXiv:1411.5394. Retrieved from <https://arxiv.org/abs/1411.5394>.
- [226] Michael Nebeling, Alexander Huber, David Ott, and Moira C. Norrie. 2014. Web on the wall reloaded: Implementation, replication and refinement of user-defined interaction sets. In *Proceedings of the 9th ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 15–24.
- [227] Matei Negulescu, Jaime Ruiz, and Edward Lank. 2012. A recognition safety net: Bi-level threshold recognition for mobile motion gestures. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, 147–150.
- [228] Jakob Nielsen. 1992. The usability engineering life cycle. *Computer* 25, 3 (1992), 12–22.
- [229] Michael Nielsen, Moritz Störing, Thomas B. Moeslund, and Erik Granum. 2004. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In *International Gesture Workshop*. Springer, Berlin, Heidelberg, 409–420. DOI: https://doi.org/10.1007/978-3-540-24598-8_38

- [230] Gerrit Niezen and Gerhard P. Hancke. 2008. Gesture recognition as ubiquitous input for mobile phones. In *Proceedings of the International Workshop on Devices That Alter Perception*, Vol. 2008. 17–21.
- [231] Laurence Nigay. 2004. Design space for multimodal interaction. In *Proceedings of the Building the Information Society*. Springer, 403–408.
- [232] Laurence Nigay and Joëlle Coutaz. 1995. A generic platform for addressing the multimodal challenge. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 98–105.
- [233] Don Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.
- [234] Donald A. Norman and Jakob Nielsen. 2010. Gestural interfaces: A step backward in usability. *Interactions* 17, 5 (2010), 46–49. DOI: <https://doi.org/10.1145/1836216.1836228>
- [235] Chris North, Tim Dwyer, Bongshin Lee, Danyel Fisher, Petra Isenberg, George Robertson, and Kori Inkpen. 2009. Understanding multi-touch manipulation for surface computing. In *Proceedings of the International Conference on Human-Computer Interaction: Part II*. 236–249.
- [236] Juliet Norton, Chadwick A. Wingrave, and J. La Viola. 2010. Exploring strategies and guidelines for developing full body video game interfaces. In *Proceedings of the International Conference on the Foundations of Digital Games*. 155–162.
- [237] Mohammad Obaid, Markus Häring, Felix Kistler, René Bühling, and Elisabeth André. 2012. User-defined body gestures for navigational control of a humanoid robot. In *Proceedings of the International Conference on Social Robotics*. 367–377.
- [238] Mohammad Obaid, Felix Kistler, Markus Häring, René Bühling, and Elisabeth André. 2014. A framework for user-defined body gestures to control a humanoid robot. *International Journal of Social Robotics* 6, 3 (2014), 383–396. DOI: <https://doi.org/10.1007/s12369-014-0233-3>
- [239] International Standards Office. 2007. *Principles and Requirements for Physical Input Devices*. Vol. ISO 9241-400. ISO.
- [240] Uran Oh and Leah Findlater. 2014. Design of and subjective response to on-body input for people with visual impairments. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, 115–122.
- [241] Eshed Ohn-Bar, Cuong Tran, and Mohan Trivedi. 2012. Hand gesture-based visual user interface for infotainment. In *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, 111–115.
- [242] Tom Ouyang and Yang Li. 2012. Bootstrapping personal gesture shortcuts with the wisdom of the crowd and handwriting recognition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2895–2904.
- [243] Oxford Dictionary. (2016) Accessed 2016. *Complexity*. Retrieved from <http://www.oxforddictionaries.com/>.
- [244] Wonkya Park and Sung H. Han. 2013. Intuitive multi-touch gestures for mobile web browsers. *Interacting with Computers* 25, 5 (2013), 335–350. DOI: <https://doi.org/10.1093/iwc/iws026>
- [245] Aman Parnami, Apurva Gupta, Gabriel Reyes, Ramik Sadana, Yang Li, and Gregory D. Abowd. 2016. Mogeste: A mobile tool for in-situ motion gesture design. In *Proceedings of the Indian Conference on Human Computer Interaction*. ACM, 35–43.
- [246] Anna Pereira, Juan P. Wachs, Kunwoo Park, and David Rempel. 2014. A user-developed 3-D hand gesture language for human-computer interaction. *The Journal of the Human Factors and Ergonomics Society* 57, 4 (2014), 607–621. DOI: <https://doi.org/dx.doi.org/10.1177/0018720814559307>
- [247] Jorge-Luis Pérez-Medina, Santiago Villarreal, and Jean Vanderdonck. 2020. A gesture elicitation study of nose-based gestures. *Sensors* 20, 24 (2020), 7118.
- [248] David N. Perkins and Gavriel Salomon. 1992. Transfer of learning. *International Encyclopedia of Education* 2 (1992), 6452–6457.
- [249] Ekaterina Peshkova, Martin Hitz, and David Ahlström. 2016. Exploring user-defined gestures and voice commands to control an unmanned aerial vehicle. In *Proceedings of the International Conference on Intelligent Technologies for Interactive Entertainment*. 47–62.
- [250] Max Pfeiffer, Dagmar Kern, Johannes Schöning, Tanja Döring, Antonio Krüger, and Albrecht Schmidt. 2010. A multi-touch enabled steering wheel: Exploring the design space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*. 3355–3360. DOI: <https://doi.org/10.1145/1753846.1753984>
- [251] Ken Pfeuffer, Jason Alexander, Ming Ki Chong, Yanxia Zhang, and Hans Gellersen. 2015. Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 373–383. DOI: <https://doi.org/10.1145/2807442.2807460>
- [252] Bastian Pflöging, Stefan Schneegass, and Albrecht Schmidt. 2012. Multimodal interaction in the car: Combining speech and gestures on the steering wheel. In *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 155–162.

- [253] Anne Marie Piper, Ross Campbell, and James D. Hollan. 2010. Exploring the accessibility and appeal of surface computing for older adult health care support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 907–916.
- [254] Pramod Kumar Pisharady and Martin Saerbeck. 2015. Recent methods and databases in vision-based hand gesture recognition: A review. *Computer Vision and Image Understanding* 141 (2015), 152–165.
- [255] Thammathip Piumsomboon, Adrian Clark, Mark Billingham, and Andy Cockburn. 2013. User-defined gestures for augmented reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*. 955–960. DOI : <https://doi.org/10.1145/2468356.2468527>
- [256] Henning Pohl, Andreea Muresan, and Kasper Hornbæk. 2019. Charting subtle interaction in the hci literature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1–15.
- [257] Prashan Premaratne and Q. Nguyen. 2007. Consumer electronics control system based on hand gesture moment invariants. *IET Computer Vision* 1, 1 (2007), 35–41. DOI : <https://doi.org/10.1049/iet-cvi:20060198>
- [258] Dmitry Pyryeskin, Mark Hancock, and Jesse Hoey. 2012. Comparing elicited gestures to designer-created gestures for selection above a multitouch surface. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 1–10.
- [259] Tsele Rakubutu, Helene Gelderblom, and Jason Cohen. 2014. Participatory design of touch gestures for informational search on a tablet device. In *Proceedings of the Southern African Institute for Computer Scientist and Information Technologists Annual Conference*.
- [260] Jef Raskin. 1994. Intuitive equals familiar. *Communications of the ACM* 37, 9 (1994), 17–19. DOI : <https://doi.org/10.1145/182987.584629>
- [261] Hanae Rateau, Laurent Grisoni, and Bruno De Araujo. 2014. Mimetic interaction spaces: Controlling distant displays in pervasive environments. In *Proceedings of the International Conference on Intelligent User Interfaces*. 89–94.
- [262] Siddharth S. Rautaray and Agrawal Anupam. 2015. Vision based hand gesture recognition for human computer interaction: A survey. *Artificial Intelligence Review* 43, 1 (2015), 1–54.
- [263] Sherri A. Rehfeld, Florian G. Jentsch, and Theresa N. Rodriguez. 2004. Memory recall for international gestures. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 48. 2604–2607.
- [264] Matthias Rehm, Nikolaus Bee, and Elisabeth André. 2008. Wave like an egyptian: Accelerometer based gesture recognition for culture specific interactions. *Proceedings of the British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction* 1 (2008), 13–22.
- [265] Yosra Rekik, Radu-Daniel Vatavu, and Laurent Grisoni. 2014. Match-up and conquer: A two-step technique for recognizing unconstrained bimanual and multi-finger touch input. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. 201–208.
- [266] Yosra Rekik, Radu-Daniel Vatavu, and Laurent Grisoni. 2014. Understanding users’ perceived difficulty of multi-touch gesture articulation. In *Proceedings of the International Conference on Multimodal Interaction*. 232–239.
- [267] David Rempel, Matt J. Camilleri, and David L. Lee. 2014. The design of hand gestures for human-computer interaction: Lessons from sign language interpreters. *International Journal of Human Computer Studies* 72, 10–11 (2014), 728–735. DOI : <https://doi.org/10.1016/J.ijhcs.2014.05.003>
- [268] Andrew D. Rice and Jonathan W. Lartigue. 2014. Touch-level model (TLM): Evolving KLM-GOMS for touchscreen and mobile devices. In *Proceedings of the ACM Southeast Regional Conference*, 53. DOI : <https://doi.org/10.1145/2638404.2638532>
- [269] Julie Rico and Stephen Brewster. 2010. Usable gestures for mobile interfaces: Evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 887–896.
- [270] Gerhard Rigoll, Andreas Kosmala, and Stefan Eickeler. 1997. High performance real-time gesture recognition using hidden Markov models. In *Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction*. 69–80.
- [271] Jessica Roberts, Leilah Lyons, Francesco Cafaro, and Rebecca Eydt. 2014. Interpreting data from within: Supporting human data interaction in museum exhibits through perspective taking. In *Proceedings of the Conference on Interaction Design and Children*. 7–16.
- [272] Suzanne Robertson and James Robertson. 2012. *Mastering the requirements process: Getting requirements right*. Addison-Wesley.
- [273] Nicla Rossini. 2003. The analysis of gesture: Establishing a set of parameters. *International Gesture Workshop* (2003), 124–131. DOI : https://doi.org/10.1007/978-3-540-24598-8_12
- [274] Nathaniel Rossol, Irene Cheng, Rui Shen, and Anup Basu. 2014. Touchfree medical interfaces. In *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, New York, NY, 6597–6600.
- [275] Gustavo Rovelo, Donald Degraen, Davy Vanacken, Kris Luyten, and Karin Coninx. 2015. Gestu-Wan-An intelligible mid-air gesture guidance system for walk-up-and-use displays. In *IFIP Conference on Human-Computer Interaction*. 368–386. DOI : https://doi.org/10.1007/978-3-319-22668-2_28

- [276] Gustavo Alberto Rovelo Ruiz, Davy Vanacken, Kris Luyten, Francisco Abad, and Emilio Camahort. 2014. Multi-viewer gesture-based interaction for omni-directional video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 4077–4086.
- [277] Quentin Roy, Sylvain Malacria, Yves Guiard, Eric Lecolinet, and James Eagan. 2013. Augmented letters: Mnemonic gesture-based shortcuts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2325–2328.
- [278] Dean Harris Rubine. 1992. The automatic recognition of gestures. *Ph.D. Dissertation*. UMI Order No. GAX92-16029 (1992).
- [279] Francisco Martinez Ruiz and Santiago Villarreal Narvaez. 2021. Eliciting user-defined zenithal gestures for privacy preferences. In *Proceedings of the International Conference on Human-Computer Interaction Theory and Applications*.
- [280] Jaime Ruiz and Yang Li. 2011. Doubleflip: A motion gesture delimiter for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2717–2720.
- [281] Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-defined motion gestures for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 197–206.
- [282] Jaime Ruiz and Daniel Vogel. 2015. Soft-constraints to reduce legacy and performance bias to elicit whole-body gestures with low arm fatigue. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 3347–3350.
- [283] Napa Sae-Bae, Kowsar Ahmed, Katherine Isbister, and Nasir Memon. 2012. Biometric-rich gestures: A novel approach to authentication on multi-touch devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 977–986.
- [284] Napa Sae-Bae, Nasir Memon, Katherine Isbister, and Kowsar Ahmed. 2014. Multitouch gesture-based authentication. In *Proceedings of the IEEE Transactions on Information Forensics and Security*, Vol. 9. IEEE, New York, NY, 568–582. DOI: <https://doi.org/10.1109/TIFS.2014.2302582>
- [285] Yasamin Sahaf, Narayanan Chatapuram Krishnan, and Diane J. Cook. 2011. Defining the complexity of an activity. In *Proceedings of the AAAI Workshop on Activity Context Representation Techniques and Languages*. 63–67.
- [286] Michael Schmidt and Gerhard Weber. 2014. Prediction of multi-touch gestures during input. In *Proceedings of the International Conference on Human-Computer Interaction*. Springer, New York, NY, 158–169.
- [287] Ben Schneiderman. 1998. *Designing the User Interface*. Addison-Wesley.
- [288] Marcos Serrano, Barrett M. Ens, and Pourang P. Irani. 2014. Exploring the use of hand-to-face input for interacting with head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3181–3190.
- [289] Amanda Mindy Seto. 2012. *Designing discoverable digital tabletop menus for public settings*. Master’s thesis, University of Waterloo.
- [290] Caifeng Shan, Tieniu Tan, and Yucheng Wei. 2007. Real-time hand tracking using a mean shift embedded particle filter. *Pattern Recognition* 40, 7 (2007), 1958–1970. DOI: <https://doi.org/10.1016/J.patcog.2006.12.012>
- [291] Shaikh Shawon Arefin Shimon, Sarah Morrison-Smith, Noah John, Ghazal Fahimi, and Jaime Ruiz. 2015. Exploring user-defined back-of-device gestures for mobile devices. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. 227–232.
- [292] Beat Signer, Ueli Kurmann, and Moira Norrie. 2007. iGesture: A general gesture recognition framework. In *Proceedings of the 9th International Conference on Document Analysis and Recognition*, Vol. 2. IEEE, New York, NY, 954–958.
- [293] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multi-modal feedback in motor learning: A review. *Psychonomic Bulletin and Review* 20, 1 (2013), 21–53.
- [294] Chaklam Silpasuwanchai and Xiangshi Ren. 2014. Jump and shoot!: Prioritizing primary and alternative body gestures for intense gameplay. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 951–954.
- [295] Chaklam Silpasuwanchai and Xiangshi Ren. 2015. Designing concurrent full-body gestures for intense gameplay. In *Proceedings of the International Journal of Human-Computer Studies*, Vol. 80. 1–13. DOI: <https://doi.org/10.1016/J.ijhcs.2015.02.010>
- [296] Lucio Davide Spano, Antonio Cisternino, Fabio Paternò, and Gianni Fenu. 2013. GestIT: A declarative and compositional framework for multiplatform gesture definition. In *Proceedings of the SIGCHI Symposium on Engineering Interactive Computing Systems*. ACM, 187–196.
- [297] Maximilian Speicher and Michael Nebeling. 2018. Gesturewiz: A human-powered gesture design environment for user interface prototypes. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–11.
- [298] Quentin Stafford-Fraser and Peter Robinson. 1996. Brightboard: A video-augmented environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 134–141.
- [299] Helman Stern, Juan Wachs, and Yael Edan. 2007. A method for selection of optimal hand gesture vocabularies. In *Proceedings of the International Gesture Workshop*. Springer, 57–68.

- [300] Helman I. Stern, Juan P. Wachs, and Yael Edan. 2008. Designing hand gesture vocabularies for natural interaction by combining psycho-physiological and recognition factors. *International Journal of Semantic Computing* 2, 1 (2008), 137–160. DOI: <https://doi.org/10.1142/S1793351X08000385>
- [301] Christian Stöbel and Lucienne Blessing. 2010. Mobile device interaction gestures for older users. In *Proceedings of the Nordic Conference on Human-Computer Interaction: Extending Boundaries*. 793–796.
- [302] Christian Stöbel and Lucienne Blessing. 2010. Tap, swipe and pinch: Designing suitable multi-touch gestures for older users. In *Proceedings of the International Design Conference*. 463–472.
- [303] Christian Stöbel, Hartmut Wandke, and Lucienne Blessing. 2009. An evaluation of finger-gesture interaction on mobile devices for elderly users. *Prospektive Gestaltung Von Mensch-Technik-Interaktion* 8 (2009), 470–475.
- [304] Christian Stöbel, Hartmut Wandke, and Lucienne Blessing. 2009. Gestural interfaces for elderly users: Help or hindrance? In *Proceedings of the International Conference on Gesture in Embodied Communication and Human-Computer Interaction*. 269–280.
- [305] Anselm Strauss and Juliet Corbin. 1990. *Basics of Qualitative Research*. Sage.
- [306] David J. Sturman and David Zeltzer. 1993. A design method for whole-hand human-computer interaction. *ACM Transactions on Information Systems* 11, 3 (July 1993), 219–238. DOI: <https://doi.org/10.1145/159161.159159>
- [307] Jesus Suarez and Robin R. Murphy. 2012. Hand gesture recognition with depth images: A review. In *Proceedings of the 2012 IEEE RO-MAN: The 21st IEEE international Symposium on Robot and Human Interactive Communication*. IEEE, New York, NY, 411–417.
- [308] Alistair Sutcliffe. 1998. Scenario-based requirements analysis. *Requirements Engineering* 3, 1 (1998), 48–65.
- [309] Ivan E. Sutherland. 1964. Sketchpad a man-machine graphical communication system. *Simulation* 2, 5 (1964), R–3.
- [310] Feng Tian, Fei Lu, Yingying Jiang, Xiaolong Zhang, Xiang Cao, Guozhong Dai, and Hongan Wang. 2013. An exploration of pen tail gestures for interactions. *International Journal of Human-Computer Studies* 71, 5 (2013), 551–569. DOI: <https://doi.org/10.1016/j.ijhcs.2012.12.004>
- [311] Jessica J. Tran, Shari Trewin, Calvin Swart, Bonnie E. John, and John C. Thomas. 2013. Exploring pinch and spread gestures on mobile devices. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. 151–160.
- [312] Giovanni Maria Troiano, Esben Warming Pedersen, and Kasper Hornbæk. 2014. User-defined gestures for elastic, deformable displays. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. ACM, 1–8.
- [313] Khai N. Truong, Gillian R. Hayes, and Gregory D. Abowd. 2006. Storyboarding: An empirical determination of best practices and effective guidelines. In *Proceedings of the Conference on Designing Interactive Systems*. ACM, 12–21.
- [314] Edward Tse, Chia Shen, Saul Greenberg, and Clifton Forlines. 2006. Enabling interaction with single user applications through speech and gestures on a multi-user tabletop. In *Proceedings of the Working Conference on Advanced Visual Interfaces*. 336–343.
- [315] Huawei Tu, Xiangshi Ren, and Shumin Zhai. 2012. A comparative evaluation of finger and pen stroke gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1287–1296.
- [316] Huawei Tu, Xiangshi Ren, and Shumin Zhai. 2015. Differences and similarities between finger and pen stroke gestures on stationary and mobile devices. *ACM Transactions on Computer-Human Interaction* 22, 5 (2015), 5.
- [317] Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, Silvia Chyou, Jhe-Wei Lin, Pei-Jung Wu, Andries Valstar, and Mike Y. Chen. 2015. User-defined game input for smart glasses in public space. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, 3327–3336.
- [318] Jacqueline Urakami. 2012. Developing and testing a human-based gesture vocabulary for tabletop systems. *Journal of the Human Factors and Ergonomics Society* 54, 4 (2012), 636–653. DOI: <https://doi.org/dx.doi.org/10.1177/0018720811433052>
- [319] Consuelo Valdes, Diana Eastman, Casey Grote, Shantanu Thatte, Orit Shaer, Ali Mazalek, Brygg Ullmer, and Miriam K. Konkel. 2014. Exploring the design space of gestural interaction with active tokens through user-defined gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 4107–4116.
- [320] Andries van Dam. 1997. Post-WIMP user interfaces. *Communications of the ACM* 40, 2 (1997), 63–67. DOI: <https://doi.org/10.1145/253671.253708>
- [321] Radu-Daniel Vatavu. 2011. The effect of sampling rate on the performance of template-based gesture recognizers. In *Proceedings of the International Conference on Multimodal Interfaces*. 271–278.
- [322] Radu-Daniel Vatavu. 2012. Nomadic gestures: A technique for reusing gesture commands for frequent ambient interactions. *Journal of Ambient Intelligence and Smart Environments* 4, 2 (2012), 79–93.
- [323] Radu-Daniel Vatavu. 2012. User-defined gestures for free-hand TV control. In *Proceedings of the European Conference on Interactive TV and Video (Euroitv'12)*. 45–48.
- [324] Radu-Daniel Vatavu. 2013. A comparative study of user-defined handheld vs. freehand gestures for home entertainment environments. *Journal of Ambient Intelligence and Smart Environments* 5, 2 (2013), 187–211. DOI: <https://doi.org/dx.doi.org/10.3233/AIS-130200>

- [325] Radu-Daniel Vatavu. 2017. Characterizing gesture knowledge transfer across multiple contexts of use. *Journal on Multimodal User Interfaces* 11, 4 (2017), 301–314.
- [326] Radu-Daniel Vatavu. 2019. The dissimilarity-consensus approach to agreement analysis in gesture elicitation studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–13. DOI: <https://doi.org/10.1145/3290605.3300454>
- [327] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2014. Gesture heatmaps: Understanding gesture performance with colorful visualizations. In *Proceedings of the International Conference on Multimodal Interaction*. 172–179.
- [328] Radu-Daniel Vatavu, Gabriel Cramariuc, and Doina Maria Schipor. 2015. Touch interaction for children aged 3 to 6 years. *International Journal of Human-Computer Studies* 74 (2015), 54–76. DOI: <https://doi.org/10.1016/j.ijhcs.2014.10.007>
- [329] Radu-Daniel Vatavu and Matei Mancas. 2013. Interactive TV potpourris: An overview of designing multi-screen TV installations for home entertainment. In *Proceedings of the International Conference on Intelligent Technologies for Interactive Entertainment*.
- [330] Radu-Daniel Vatavu, Daniel Vogel, Géry Casiez, and Laurent Grisoni. 2011. Estimating the perceived difficulty of pen gestures. In *Proceedings of the IFIP Conference on Human-Computer Interaction*. 89–106.
- [331] 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*. ACM, New York, NY, 1325–1334. DOI: <https://doi.org/10.1145/2702123.2702223>
- [332] 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*. 1325–1334. DOI: <https://doi.org/10.1145/2702123.2702223>
- [333] 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*. 131–138.
- [334] 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 Conference on Designing Interactive Systems*.
- [335] Dong-Bach Vo, Eric Lecolinet, and Yves Guiard. 2014. Belly gestures: Body centric gestures on the abdomen. In *Proceedings of the Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. 687–696.
- [336] Daniel Vogel and Ravin Balakrishnan. 2010. Occlusion-aware interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 263–272.
- [337] Daniel Vogel and Patrick Baudisch. 2007. Shift: A technique for operating pen-based interfaces using touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 657–666.
- [338] Daniel Vogel and Géry Casiez. 2012. Hand occlusion on a multi-touch tabletop. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2307–2316.
- [339] Daniel Vogel, Matthew Cudmore, Géry Casiez, Ravin Balakrishnan, and Liam Keliher. 2009. Hand occlusion with tablet-sized direct pen input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 557–566.
- [340] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. 2018. Gesture elicitation studies for mid-air interaction: A review. *Multimodal Technologies and Interaction* 2, 4 (2018), 65.
- [341] Tijana Vuletic, Alex Duffy, Chris McTeague, Laura Hay, Ross Brisco, Gerard Campbell, and Madeleine Grealy. 2021. A novel user-based gesture vocabulary for conceptual design. *International Journal of Human-Computer Studies* 150 (2021), 102609.
- [342] Juan P. Wachs, Mathias Kölsch, Helman Stern, and Yael Edan. 2011. Vision-based hand-gesture applications. *Communications of the ACM* 54, 2 (2011), 60–71. DOI: <https://doi.org/10.1145/1897816.1897838>
- [343] Juan P. Wachs, Helman I. Stern, Yael Edan, Michael Gillam, Jon Handler, Craig Feied, and Mark Smith. 2008. A gesture-based tool for sterile browsing of radiology images. *Journal of the American Medical Informatics Association* 15, 3 (2008), 321–323. DOI: <https://doi.org/10.1197/jamia.M241>
- [344] Jagoda Walny, Bongshin Lee, Paul Johns, Nathalie Henry Riche, and Sheelagh Cpendale. 2012. Understanding pen and touch interaction for data exploration on interactive whiteboards. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2779–2788. DOI: <https://doi.org/10.1109/TVCG.2012.275>
- [345] Robert Walter, Gilles Bailly, and Jörg Müller. 2013. Strikeapose: Revealing mid-air gestures on public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 841–850.
- [346] De-Xin Wang, Qing-Bao Liu, and Mao-Jun Zhang. 2011. A multi-touch platform based on four corner cameras and methods for accurately locating contact points. *Multimedia Tools and Applications* 51, 3 (2011), 963–982. DOI: <https://doi.org/10.1007/s11042-009-0425-2>
- [347] De-Xin Wang, Zhi-Hui Xiong, and Mao-Jun Zhang. 2012. An application oriented and shape feature based multi-touch gesture description and recognition method. *Multimedia Tools and Applications* 58, 3 (2012), 497–519. DOI: <https://doi.org/10.1007/S11042-011-0730-4>

- [348] Diane Watson, Mark Hancock, Regan L. Mandryk, and Max Birk. 2013. Deconstructing the touch experience. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. 199–208.
- [349] Christian Weichel, Manfred Lau, David Kim, Nicolas Villar, and Hans W. Gellersen. 2014. Mixfab: A mixed-reality environment for personal fabrication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3855–3864.
- [350] Martin Weigel, Vikram Mehta, and Jürgen Steimle. 2014. More than touch: Understanding how people use skin as an input surface for mobile computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 179–188.
- [351] Wayne Westerman, John G. Elias, and Alan Hedge. 2001. Multi-touch: A new tactile 2-d gesture interface for human-computer interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 45. 632–636. DOI: <https://doi.org/10.1177/154193120104500612>
- [352] Daniel Wickerroth, Paul Benölken, and Ulrich Lang. 2009. Manipulating 3D content using gestures in design review scenarios. *International Journal of Information 1*, 4 (2009), 243.
- [353] Karl E. Wieggers. 1999. First things first: Prioritizing requirements. *Software Development 7*, 9 (1999), 48–53.
- [354] Daniel Wigdor, Hrvoje Benko, John Pella, Jarrod Lombardo, and Sarah Williams. 2011. Rock and rails: Extending multi-touch interactions with shape gestures to enable precise spatial manipulations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1581–1590.
- [355] Daniel Wigdor, Clifton Forlines, Patrick Baudisch, John Barnwell, and Chia Shen. 2007. Lucid touch: A see-through mobile device. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 269–278.
- [356] Daniel Wigdor, Sarah Williams, Michael Cronin, Robert Levy, Katie White, Maxim Mazeev, and Hrvoje Benko. 2009. Ripples: Utilizing per-contact visualizations to improve user interaction with touch displays. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 3–12.
- [357] Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture* (1st. ed.). Morgan Kaufmann Publishers Inc.
- [358] Julie R. Williamson. 2012. *User Experience, Performance, and Social Acceptability: Usable Multimodal Mobile Interaction*. Ph.D. Dissertation. University of Glasgow.
- [359] Andrew D. Wilson. 2005. Playanywhere: A compact interactive tabletop projection-vision system. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 83–92.
- [360] Andrew D. Wilson, Shahram Izadi, Otmar Hilliges, Armando Garcia-Mendoza, and David Kirk. 2008. Bringing physics to the surface. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 67–76.
- [361] Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the guessability of symbolic input. In *Proceedings of the ACM Conference on Human Factors in Computing Systems Extended Abstracts*. 1869–1872.
- [362] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-Defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1083–1092.
- [363] Jacob O. Wobbrock, Brad A. Myers, and Htet Htet Aung. 2008. The performance of hand postures in front-and back-of-device interaction for mobile computing. *International Journal of Human-Computer Studies 66*, 12 (2008), 857–875.
- [364] Jacob O. Wobbrock, Brad A. Myers, and John A. Kembel. 2003. EdgeWrite: A stylus-based text entry method designed for high accuracy and stability of motion. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 61–70.
- [365] Jacob O. Wobbrock, Andrew D. Wilson, and Yang Li. 2007. Gestures without libraries, toolkits or training: A \$1 recognizer for user interface prototypes. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 159–168.
- [366] Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. 559–575. Taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-dependent requirements. In *Proceedings of the International Conference on Human-Computer Interaction - Volume Part I*. DOI: https://doi.org/10.1007/978-3-642-23774-4_45
- [367] Katrin Wolf, Robert Schleicher, and Michael Rohs. 2014. Ergonomic characteristics of gestures for front- and back-of-tablets interaction with grasping hands. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. 453–458. DOI: <https://doi.org/10.1145/2628363.2634214>
- [368] Michael Wright, Chun-Jung Lin, Eamonn O’Neill, Darren Cosker, and Peter Johnson. 2011. 3D gesture recognition: An evaluation of user and system performance. In *Proceedings of the International Conference on Pervasive Computing*. 294–313.
- [369] Michael A. E. Wright. 2015. *Supporting the Transfer of Learning of Freehand Gestures*. Ph.D. Dissertation. University of Bath.
- [370] Huiyue Wu, Jiayi Liu, Jiali Qiu, and Xiaolong Luke Zhang. 2019. Seeking common ground while reserving differences in gesture elicitation studies. *Multimedia Tools and Applications 78*, 11 (2019), 14989–15010.

- [371] Huiyue Wu and Jianmin Wang. 2012. User-Defined body gestures for TV-Based applications. In *Proceedings of the International Conference on Digital Home*. 415–420.
- [372] Huiyue Wu, Jianmin Wang, and Xiaolong Luke Zhang. 2016. User-centered gesture development in TV viewing environment. *Multimedia Tools and Applications* 75, 2 (2016), 733–760. DOI : <https://doi.org/10.1007/s11042-014-2323-5>
- [373] Huiyue Wu, Yu Wang, Jiayi Liu, Jiali Qiu, and Xiaolong Luke Zhang. 2020. User-defined gesture interaction for in-vehicle information systems. *Multimedia Tools and Applications* 79, 1 (2020), 263–288.
- [374] Mike Wu and Ravin Balakrishnan. 2003. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 193–202.
- [375] Mike Wu, Chia Shen, Kathy Ryall, Clifton Forlines, and Ravin Balakrishnan. 2006. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *Proceedings of the IEEE International Workshop on Horizontal Interactive Human-Computer Systems*. IEEE, New York, NY, 185–192. DOI : <https://doi.org/10.1109/TABLETOP.2006.19>
- [376] Haijun Xia, Bruno Araujo, Tovi Grossman, and Daniel Wigdor. 2016. Object-oriented drawing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 4610–4621.
- [377] Haijun Xia, Tovi Grossman, and George Fitzmaurice. 2015. NanoStylus: Enhancing input on ultra-small displays with a finger-mounted stylus. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 447–456.
- [378] Haijun Xia, Ken Hinckley, Michel Pahud, Xiao Tu, and Bill Buxton. 2017. WritLarge: Ink unleashed by unified scope, action, and zoom. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3227–3240.
- [379] Koji Yatani, Kurt Partridge, Marshall Bern, and Mark W. Newman. 2008. Escape: A target selection technique using visually-cued gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 285–294.
- [380] Hui Ye, Kin Chung Kwan, Wanchao Su, and Hongbo Fu. 2020. ARAnimator: In-situ character animation in mobile AR with user-defined motion gestures. *ACM Transactions on Graphics* 39, 4 (2020), 83–1.
- [381] Wendy Yee. 2009. Potential limitations of multi-touch gesture vocabulary: Differentiation, adoption, fatigue. In *Proceedings of the International Conference on Human-Computer Interaction. Part II: Novel Interaction Methods and Techniques*. 291–300.
- [382] Dongwook Yoon, Ken Hinckley, Hrvoje Benko, François Guimbretière, Pourang Irani, Michel Pahud, and Marcel Gavrilu. 2015. Sensing tablet grasp + micro-mobility for active reading. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. 477–487.
- [383] Ionuț-Alexandru Zaiți, Ștefan-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 (2015), 821–838.
- [384] Bruno Zamborlin, Frederic Bevilacqua, Marco Gillies, and Mark D’Inverno. 2014. Fluid gesture interaction design: Applications of continuous recognition for the design of modern gestural interfaces. *ACM Transactions on Interactive, Intelligent Systems* 3, 4 (2014), 22. DOI : <https://doi.org/10.1145/2543921>
- [385] Adam Zarek, Daniel Wigdor, and Karan Singh. 2012. SNOUT: One-handed use of capacitive touch devices. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. 140–147.
- [386] Robert Zeleznik and Timothy Miller. 2006. Fluid inking: Augmenting the medium of free-form inking with gestures. In *Proceedings of the Graphics Interface*. Canadian Human-Computer Communications Society, 155–162.
- [387] Shumin Zhai and Per Ola Kristensson. 2012. The word-gesture keyboard: Reimagining keyboard interaction. *Communications of the ACM* 55, 9 (2012), 91–101. DOI : <https://doi.org/10.1145/2330667.2330689>
- [388] Xiaoyi Zhang, Harish Kulkarni, and Meredith Ringel Morris. 2017. Smartphone-based gaze gesture communication for people with motor disabilities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2878–2889.
- [389] Rui Zhao. 1993. Incremental recognition in gesture-based and syntax-directed diagram editors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 95–100.
- [390] Kening Zhu, Ma Xiaojuan, Chen Haoyuan, and Liang Miaoyin. 2017. Tripartite effects: Exploring users’ mental model of mobile gestures under the influence of operation, handheld posture, and interaction space. *International Journal of Human-Computer Interaction* 33, 6 (2017), 443–459.

Received August 2020; revised November 2021; accepted December 2021