Eliciting User-Defined Touch and Mid-air Gestures for Co-located Mobile Gaming

CHLOE NG and NICOLAI MARQUARDT, University College London, UK

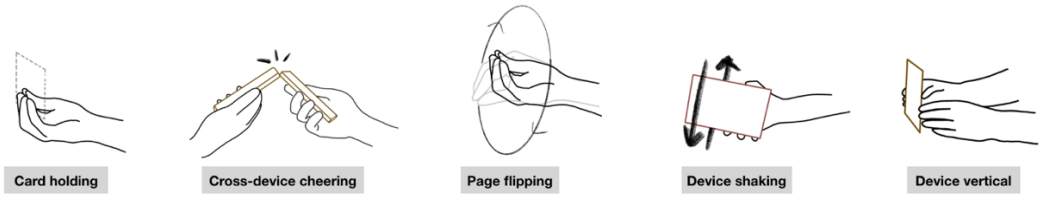


Fig. 1. Example mid-air gestures from our elicitation study of common board- and card-game tasks.

Many interaction techniques have been developed to best support mobile gaming – but developed gestures and techniques might not always match user behaviour or preferences. To inform this design space of gesture input for co-located mobile gaming, we present insights from a gesture elicitation user study for touch and mid-air input, specifically focusing on board and card games due to the materiality of game artefacts and rich interaction between players. We obtained touch and mid-air gesture proposals for 11 game tasks with 12 dyads and gained insights into user preferences. We contribute our classification and analysis of 622 elicited gestures (showing more collaborative gestures in the mid-air modality), resulting in a consensus gesture set, and agreement rates showing higher consensus for touch gestures. Furthermore, we identified interaction patterns – such as benefits of situational awareness, social etiquette, gestures fostering interaction between players, and roles of gestures providing fun, excitement, and suspense to the game – which can inform future games and gesture design.

CCS Concepts: • **Human-centered computing → Empirical studies in interaction design.**

Additional Key Words and Phrases: gesture elicitation, hand gestures, mobile device, tabletop gaming.

ACM Reference format:
Chloe Ng and Nicolai Marquardt. 2022. Eliciting User-Defined Touch and Mid-air Gestures for Co-located Mobile Gaming. Proc. ACM Hum.-Comput. Interact. 6, ISS, Article 569 (December 2022), 25 pages. https://doi.org/10.1145/3567722

1 INTRODUCTION

To allow for more engaging and entertaining experiences in mobile devices, many research and commercial projects explored creative forms of digital input aside from using touch input on capacitive screens [2,8,39,43]. For example, some commercial games use alternative input techniques such as tilting and shaking (enabled by built-in gyroscope and accelerometer) for engaging gameplay [14,16,22]. Most existing products and research focus on single-player and online multiplayer gameplay, and a large body of research for co-located mobile gaming focuses in particular on game mechanic design and strategies [4,5,18,20].

Authors' address: University College London, Gower Street, WC1E 6BT, London, United Kingdom

2573-0142/2022/12-ART569 $15.00
https://doi.org/10.1145/3567722

To inform the overall design space of gestural interaction in co-located mobile multi-player gaming, we explore the use of gestures inspired by traditional physical multiplayer tabletop games such as board and card games due to their clearly defined game tasks and the materiality of the game materials cherished by players [33]. This is inspired by and extends beyond insights from studies around digital tabletop games, where research explored how to provide rich game interaction for engaging gameplay – for example, by incorporating tangibles [3], multimodality and face-to-face interaction [10], VR-board-game hybrids [41], or multi-device gameplay [51]. Our work also relates to studies done about gestural interaction in mobile gaming, such as Skov et al.'s comparison of cross-device gestures for card games [40].

A better understanding of people's behaviours and preferences for gestures has the potential to inform future designs of such gameplay interaction techniques in mobile gaming. To address the challenge of understanding gesture use and preferences from a user perspective, we conducted a gesture elicitation study [46] to help develop user-defined gestures – similar to other elicitation studies, e.g., for surface computing [50], groupware [27], or connecting mobile phones [19]. Expanding on such more general studies of motion gestures with phones [19,35], with our study we seek to find out how results and observations made in a gesture elicitation for game tasks can inform gesture design in co-located multiplayer mobile games.

In particular, we elicited touch and mid-air gestures for eleven common board and card game tasks from twelve dyads of participants (see examples in Figure 1). Informed by previous gesture elicitation studies [9,25,35,50], we classified the elicited gestures into five dimensions and calculated the revised agreement rates [46]. We collected and analysed subjective ratings on user preference. Lastly, we applied thematic analysis [6] to analyse participants' comments, which were made both verbally during the elicitation and in written form in a post-study questionnaire, resulting in six themes. Overall, our contributions are:

1. A user-elicited mobile co-located gaming set of 662 gestures, and a consensus set of 13 touch and mid-air gestures for the eleven game tasks (including revised agreement rate of 22 referent-modality combinations).
2. A classification showing gesture distributions in five dimensions, namely nature, space, actor, temporal flow, and complexity.
3. Six themes generated from the thematic analysis of interviews, think-aloud data, and written comments: e.g., about importance of 'fun' gestures enriching gameplay, social etiquette for gestures, use of collaborative gestures, and benefits of situational awareness.
4. A discussion on how the quantitative and qualitative data together generate implications for gesture design, game design, and sensing technologies in general.

As mobile gaming encompasses a wide variety of games (loosely defined in [22]), the focus of this work is mobile games played on personal digital devices such as smartphones and tablets. We use this narrower definition to differentiate the games we are investigating from those played on dedicated game consoles or controllers for co-located gaming.

2 RELATED WORK

2.1 Interaction Techniques in Mobile Gaming

A strand of mobile game development explores alternative interaction techniques, both in the research community and the gaming industry. In the single-player domain, there are multiple efforts to make use of built-in sensors in mobile devices to implement new interaction techniques.

Way Out [43] is a single-player maze adventure game that uses the motion sensor and front-facing camera in smartphones to create a realistic game environment through around-body interactions. The player can explore game scenes by tilting and orienting smartphones, enabling navigation in three-dimensional space. ActivRunner [2] is a movement-based mobile game using the built-in gyroscope and accelerometer to detect whole-body movement. Similarly, a multiplayer example would be Road Rager, developed by Brunnberg [8], which uses proximity and geolocation of mobile devices to enable players to interact from near-by moving vehicles. New input control in commercial products also emerged in recent years, introducing interaction techniques such as tilting, throwing and finger tapping back of smartphones to enhance the mobile game experience [14,16].

Besides input modalities, research focuses on game mechanics such as game rules and tasks. TaxiTrouble [4], a social game played on tablets and smartphones was developed to enhance interaction through competitive and collaborative game tasks. Similarly, Brick is an augmented reality mobile game [5] that encourages player interaction through collaborative tricks and real-time status updates. Lautamäki and Suomela [20] explored the use of Bluetooth for proximity sensing in multiplayer games with mobile phones. Kim et al. [18] looked into the application of new spatial multi-device configuration in casual games. Although these research projects explored interaction techniques enriching the mobile gaming experience, they did not cover the use of hand gestures to enable interaction between players. Current vision-based gesture recognition approaches (for gestures in gaming) vary in terms of accuracy and performance [17]. In this context, it is important to better understand user-defined gestures for gaming that would enrich game experiences but at the same time is feasible in terms of detection accuracy and performance. As touch and mid-air gestures increasingly become part of mobile (co-located) gaming experiences, our study aims at identifying common gesture uses and interaction patterns that can inform the selection and design of such gestures.

2.2 Digital Tabletop Games

Past research has studied the use of tabletop devices for board and card games. One of the early digital board games in HCI was the STARS system [20], which contains a tabletop device, a large wall display, a handheld device, and tangible pieces, preserving the materiality of manipulating game artefacts cherished by gamers. WeatherGod [3] is a digital board game that integrates physical play pieces onto a tabletop device, retaining tangible interaction. A digital version of Monopoly, Copoly [51], used handheld devices for private information for individual player interacting with a large tabletop for a shared view for all players. Lobunets and Prinz [21] developed a digital prototype for Uno card play using handheld devices and large public display. Melissa et al. [34] derived a model that showcases the activities performed by digital tools alongside with physical components in hybrid digital boardgames.

One attempt at implementing multiplayer card games digitally is Surface-Poker on the Microsoft Surface digital table [10], making use of common poker game gestures such as shielding to hide private card set from other players on the same table, circling coins to metaphorically select multiple coins, and double tapping the screen to digitally check in poker game. Using solely a shared tabletop display did not allow easy private information retrieval, which was addressed by Shirazi et al.'s cross-device PokerSurface [38] that combined mobile devices with tabletop, along with an expert design of touch and motion gestures to facilitate card play. In the same context, Skov et al. [40] investigated six expert-designed interaction techniques for card playing using mobile devices through a usability study. A challenge in some of these systems is that larger

displays often could not provide direct manipulation of game objects. For imperfect information games, tabletop displays might also require smaller handheld devices to provide a private view. Although some argued that digital tabletop device allows co-located casual home gaming [41], it has not gained the same level of popularity as smaller mobile devices like tablets and smartphones, which have largely penetrated the fabric of everyday life. Besides this form factor issue, gestural inputs in the preceding work were often designed by experts. Resulting gestures might be hard to learn or memorised by users – a long-discussed challenge in gesture design [47]. Overall, our elicitation study aims to complement expert design-driven approaches for mobile gesture design (e.g., [40]), and further investigate gameplay gestures on mobiles rather than tabletop settings.

2.3 User Elicitation Studies for Surface Interaction

User elicitation is one form of participatory design [36] to include users' mental models and proposals in designing new interaction techniques. Elicitation studies aim to invoke easy-to-learn and memorable user-defined gestures instead of gestures that are optimised for machine recognition [28]. In an elicitation, participants are asked to propose gestures to achieve tasks (known as referents) in a specified modality. Early research in gesture elicitation studies was to explore touch-based interaction techniques for surface computing [50]. Since then, user elicitation has been applied widely in areas such as mobile devices [12,35], home environments [7,48], public display [13,32], and in-vehicle system [23]. Gesture types ranged from micro-hand gesture [9,42] to whole-body movement [32,39].

In mobile gaming, a few elicitations studied single-player game interaction. For example, Silpasuwanchai and Ren [39] established a full-body gesture set for intense gameplay with choice-based elicitation. Tung et al. [44] developed a set of user-defined touch and non-touch input for gaming with smart glasses in public space. The user elicitation work closest to co-located mobile gaming was in cross-device interaction. Here, past user elicitation studies on gestural control fall into two branches – digital object manipulation [9,31] and collaborative sensemaking [11,15,30]. Kray et al. [19] studied how mobile phones can be used to gesture in a multi-device environment with public display and tabletops. Seyed et al. [37] elicited hand gestures for transferring image across tablets, tabletops and wall display. The resulting gestures were of low convergence. In collaborative sensemaking, Desolda et al. [11] performed an elicitation with tablets and phones, resulting in spatially aware and synchronous gestures such as directional flick to transfer text and device bumping for visualising data across devices, as well as connecting devices through proximity or bumping. Dingler et al. [12] went a step further to consolidate unified gestures for active reading on phone, watch, and glasses through gesture elicitation and validation with transferability score. Our work draws on the gesture elicitation methodology to look into interaction between players and to expand cross-device interaction to mobile gaming. In particular, our specific focus on mobile co-located gaming has the potential to complement other more general gesture vocabularies – such as for cross-device transfer [37], mobiles in multi-device spaces [19], and motion input with mobiles [35] – with the goal of having a better understanding of gestures that match expectations and common practice for mobile co-located gaming.

2.4 Best Practices for Elicitation Studies

We briefly summarise best practices for conducting and evaluating gesture elicitation studies. An elicitation result depends on the participants; therefore, the resulting proposals could be affected by a participant's previous experience in technology (i.e., legacy bias). Morris et al. [26] proposed three techniques to mitigate legacy bias. *Priming* is to show new and rare interaction techniques

the elicitation so that participants are encouraged to think without being confined by existing commercial technologies. Other techniques include *production*, in which participants are asked to propose a minimum number of interaction techniques, and *partnering*, in which two or more participants propose in the same session to stimulate one another.

In terms of gesture analysis, researchers have developed metrics to capture important characteristics in gesture proposals. The most frequently adopted one is Wobbrock et al.'s agreement rate [49], which captures the degree of consensus among users. This current work uses the recently revised agreement rate by Vatavu and Wobbrock [46] through their toolkit for calculation of agreement, disagreement, co-agreement rate and statistical tests for relationships between proposals. Other's metrics include the consensus-distinct and max-consensus ratios [25], which are suitable for elicitation when the number of proposals from each participant is not fixed, meaning participants can freely propose gestures without an upper limit. Although metrics help standardise the evaluation of gesture proposals, each of them only investigates one aspect of the proposals.

To better understand characteristics of the gestures proposed as a whole, numerous studies [1,50] have attempted to classify gestures under different dimensions, ranging from gesture complexity to kinematic impulse [35], from spatial relationships [24] to the body parts involved [44]. These characterization results in the classification breakdowns could help researchers understand patterns of the gesture proposals. Our elicitation methodology in this paper is informed by those elicitation studies, and we employ a customised classification breakdown and calculation of agreement rates for the analysis which we will explain in the Results section.

3 GESTURE ELICITATION STUDY FOR TOUCH/GESTURE IN MOBILE GAMING

In this section, we explain the details of our gesture elicitation study for touch/mid-air gestures for co-located mobile gaming, focusing on tasks common in multiplayer card and board games.

3.1 Objectives

We chose general board and card game tasks for our elicitation because they are well-defined tasks that cover the interaction between players, players, and the shared space, as well as players and the game artefacts. Our objective was to elicit user-defined *Touch and Mid-air gestures*: Touch gestures are defined as a gestural interaction that contains on-screen touch, with or without motion gestures. An example of this would be "finger swiping on-screen with one hand while moving the device forward with another hand". Mid-air gestures are gestures performed in the air, with or without on-screen touch gestures or motion gestures. An example of this would be a "grab and release" gesture in which the "five fingers centralise on screen, then releasing the clenched hand in air". For simplicity, pure motion gestures in which the device itself is manipulated to perform an action in the air are also considered as mid-air gestures. An example of this would be wobbling the device.

The above definitions were motivated by observations from a small pilot study, where participants often proposed gestures that include a combination of touch, mid-air, and motion gestures. The definitions allow us to compare touch-based gestures against alternative gestures that require gesture detection with sensors other than a capacitive screen. We intend to understand the mental models and reasoning behind the gesture proposal for these common game tasks, instead of reaching a definitive set of final gestures for implementation. We examine the patterns of gesture proposals for both the gesture proposals and player interaction through quantitative metrics and qualitative data from the elicitation.

3.2 Participants

Twenty-four participants (identifying as 8 males, 16 females) were recruited in pairs of two through a university participant pool and snowball sampling. Participants aged between 20 and 55 years (M = 28.04, SD = 8.88) and worked in diverse occupations. Education levels ranged from high school to Master's graduate, with participants from Europe (n=12), Asia (n=7), North (n=3) and South (n=2) America. We recruited this group of participants with heterogeneous backgrounds in response to the overlooked issue of participant profile in similar studies [47].

All participants indicated that they played card or board games occasionally (87.5%, n=21) to frequently (12.5%, n=3). All participants used smartphones frequently. Three participants (12.5%) never used tablets while 21 (87.5%) used tablet occasionally to frequently. Nine (37.5%) had used smartwatches, and ten (41.7%) had experience using motion-capturing devices such as Microsoft Kinect and Nintendo Wii. None of the participants had any experience in micro-gestural recognition devices such as the Myo Armband [29]. Two (8.3%) participants were left-handed.

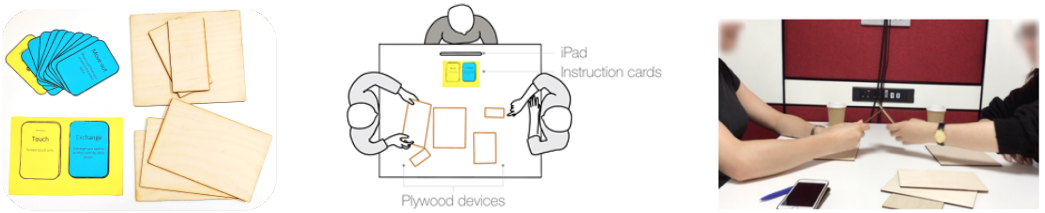


Fig. 2. (Left) plywood boards and instruction cards. (Center) top view, (Right) side view of the elicitation study setup.

3.3 Apparatus

Mock-up devices (made with plywood) were used instead of real tablets or phones (Figure 2 left) to minimise the effect on mental models from existing interaction with digital devices and disconnect participants from platform-specific gestures they might be used to. Instruction cards show referent and modality meanings and were displayed for participants' reference throughout the elicitation. Both the plywood boards and instruction card samples are shown in Figure 2 (left). Pre- and post-study questionnaires ask about the background subjective ratings and free-form feedback about the elicitation study and the gestures proposed. Elicitation sessions were video recorded with an iPad, capturing hand and arm movements from the side, and the communication between the pair of participants, and the author as shown in Figure 2 (center and right).

3.4 Referents

To consolidate the list of 11 task referents (Table 1) that were common and frequently performed in analogue tabletop games, we draw upon Skov et al.'s usability study [40] of six cross-device interaction techniques for playing and drawing cards. We extend these two game tasks by further classifying tasks into transfer between devices and pure manipulation of artefacts, resulting in the two categories Transfer and Manipulating artefacts. To specify the device type, Transfer is further broken down into two sub-categories, namely private-to-private (transfer to another player's device) and private-to-public (transfer to a shared space), as we would like to explore how the actions and the device type might affect (if any) the interaction techniques proposed. We include Manipulating artefacts because we have also considered tasks such as dice rolling and card shuffling that are common in tabletop games and are underexplored in current digital tabletop games.

Table 1: The list of 11 referents used in this study.

Category	Task	Description
Transfer (Private-Private)	1. Give card	Give a card to another player
	2. Take card	Take a card from another player's hand
	3. Exchange	Exchange cards between two players
	4. Give chip	Give chips/tokens to another player
	5. Take chip	Take chips/tokens from another player
Transfer (Private-Public)	6. Move out	Move a card from private to public space
	7. Move in	Move a card from public to private space
Artefact manipulation	8. Flip	Flip/turn over a card
	9. Shuffle cards	Shuffle a deck of cards
	10. Shuffle dice	Shuffle dice
	11. Stack	Stack chips/token

3.5 Procedure

After obtaining consent, participants fill in a pre-study questionnaire on demographics, technology usage, and general tabletop gameplay habits. To help participants think out of the box without constraints, the researcher *primed* [26] the study by providing examples of novel gestural control technologies verbally, such as the Myo Armband application in mid-air gestural control for biking [29]. Participants then chose one of the plywood boards, which served as a proxy device they use during the elicitation. All remaining boards remained on the desk – participants were told that they could interchange boards of different sizes at any time (e.g., to express ideas for different device sizes as they go). Participants were asked to think aloud while they propose at least 2 gestures for each referent-modality combination (an example of combination: *Give card + Mid-air gesture*). They were also encouraged to discuss with their partners about the proposals. We encouraged the participants to suggest gestures freely without concerns for technological feasibility.

At the start of the elicitation, the researcher randomly drew a task card (referent) from the shuffled deck of 11 referents, and then verbally described the task. The *Modality Card* was then randomly drawn and verbally explained. The dyad then proposed gestures for the given referent-modality combination. They were invited to reason for their gestures throughout the elicitation. After two or more gestures proposed and the agreement to move on, the dyad then continues to propose gestures for the same referent in the other modality. After each referent, the dyad was asked to come to a consensus for a favourite gesture for the two referent-modality combinations. They were asked follow-up questions about their decisions, and their opinions about the touch and mid-air gestures chosen favourite. The same procedures repeated for the other 20 referent-modality combinations. After the whole elicitation, we asked participants about their preference and views on both touch and mid-air gesture. The whole study lasted for around 45 minutes. The study was approved by *Anonymous Institution* Ethical Review Board.

3.6 Data Collection and Analysis

Observational notes, video recordings, post-study interview data were collected. The pre- and post-study questionnaires contain demographic, technology usage habits, Likert-scale and free-form responses. Descriptive statistics were generated from the Likert-scale questions. All gestures were coded (and used the AGATe 2.0 tool [46] for agreement rate calculation). In the interviews we discussed general opinions on the gestures chosen, their favourite gestures, and their

preference on using gestural control. The transcription and free-form written feedback were thematically analysed [6]. Our analysis process was (LA = *lead author*, AA = *all authors*): (1, LA) jotting down parts of the utterances during the elicitation session; (2, LA) watching video recordings on the same day to identify and record all gesture proposals, verbal utterances for each proposal, and the exit interview; (3, LA) labelling transcription, i.e., coding the dataset; (4, AA) generating initial themes from the codes; (5, AA) reviewing themes after coding all transcription; (6, LA) watching all video recordings again and transcribe any missing verbal utterances, while reviewing the codes and themes again; (7, AA) defining and prioritising themes, as well as identify relevant quotes; and (8, LA) expanding on the themes in written form, supported by relevant quotes. This coding and analysis were conducted by the lead author, and were reviewed and checked across the authors.

4 RESULTS

We collected a total of 662 gesture proposals, with 286 distinct gestures (by referents) from our twelve pairs of participants. Participant pairs were encouraged to propose as many gestures as they wish, with a minimum of two, resulting in 30.09 gesture proposals on average (SD=1.69) for each referent-modality combination.

Table 2: Classification of elicited gestures for multiplayer game control based on 662 gesture proposals.

Nature	*Physical*	Gesture physically acts on objects as in real life
	Metaphorical	Gesture has a metaphor for task action
	Symbolic	Gesture depicts a symbol
	Abstract	Gesture is arbitrary
Space	*In air*	Gesture performed in air
	On bezel	Gesture performed in the surroundings of the device
	On device	Gesture performed on screen
	Using device	Gesture performed using the device
	Hybrid	Gesture performed a combination of in air, on bezel, or using device.
Actor	*Single*	Only one single person performing the gesture
	Double symmetric	Two people collaborate with same gesture on both sides
	Double asymmetric	Two people collaborate with different gestures
Temporal Flow	*Discrete*	Response occurs after the gesture
	Continuous	Response occurs as the gesture occurs
Complexity	*Simple*	Gesture consists of one step
	Compound	Gesture consists of more than one step

4.1 Gesture Coding Procedure

Video recordings of the gesture elicitation were analysed in three rounds: The first round was performed on the same day as the session to minimise memory loss. The experimenter coded gestures in plain English. The second round focused on utterances, the think-aloud protocol and exit interviews. The third round served to refine and verify. During gesture coding, we grouped similar gestures together by omitting minor differences that the participants explicitly claimed it did not matter precisely, for example, a Touch gesture for Shuffle dice, one or multiple fingers "rubbing", "circulating", and "scribbling" on screen are all categorised into Circulate.

Often times, when participants struggled to propose different gestures, they would arbitrarily vary the number of fingers used with no strong rationale behind but to "make [the gesture]

different" from the other. This observation was also documented in previous hand gesture elicitation literature [9,50]. In other cases, participants often preferred using multiple fingers when manipulating multiple objects such as chips and cards but did not have a strong preference for the exact number of fingers, if they were more than two. Due to the above differences in mental models, we simplified the gesture codes by combining one and two fingers (index and middle) to reduce the randomness in arbitrary proposals, as well as combining three and four fingers for multiple object manipulation.

From the original 662 gestures, we were able to reduce the proposals into 286 distinct gestures by referent, of which 71 gestures with a distinct threshold of 2 i.e., gestures proposed for at least twice in each referent-modality combination. More details of the distinct threshold can be found in Morris et al.'s work [25].

4.2 Classifying Referent-Modality Gesture Combinations

4.2.1 Classification of game gestural inputs. A taxonomy of classification allows us to systematically present the nature of the gestures proposed and to explore any patterns in the gesture proposals. While previous literature employed a wide range of dimensions delineated in Section 2.4, we extracted and edited the most relevant ones from previous work [35,45,50] for comparability, and added a new dimension Actor to our classification. All gestures were classified under five areas – Nature, Space, Actor, Temporal Flow, and Complexity. For a summary of the classification definition, please refer to Table 2. Some dimensions are designed for specific modality or referent type only. For this reason, the resulting classification is presented in overall level and divided into modality and referent type level in the next section. For now, we will explain the classification with examples in the following.

- ***Nature:*** Nature categorizes the mental model of the gesture proposal. Physical means the gesture represents a real-life physical manipulation of the object, such as dragging a card on screen from one device to another to represent the transfer of card. This is because the participant explained this action to be sliding the card to the opposite in the physical world. Metaphorical means the gesture presents a metaphor for the referent performed, such as performing a hand-washing gesture in air to represent shuffling cards. Symbolic is a gesture that depicts a symbol, such as writing a checkmark on the screen with an index finger to represent flipping a card. Abstract includes gestures that do not fall into the above three categories or are from legacy control such as WIMP menu control.
- ***Space:*** The Space dimension is modified to capture the use of device directly to perform a mid-air interaction in Using device, such as motion gesture of placing the device perpendicular to the table. On bezel is when a gesture is performed around the device, such as palm sliding forward next to the device. Gestures performed In air can be twirling index finger for flipping a card. Hybrid is a combination of the above spatial dimensions, which is common in our dataset. On device captures all Touch gesture due to the nature of the modality.
- ***Actor:*** We extend the classification of elicited gestures by the Actor dimension, to cover the collaborative features of gestures involving two players, which have not been looked into in previous literature. Most Artefact manipulation gestures (e.g., shuffle cards/dice. See Table 1) are Single actor. Double symmetric gestures are the same gestures performed by two players, for example, a dyad shaking their devices with front facing each other simultaneously. Double asymmetric requires two players as well but

performing different gestures. For example, a dyad holding their devices simultaneously with the top of the device in contact, but only one of them swipes on the screen.

- ***Temporal:*** In Temporal flow, Discrete gesture can be shuffling dice after one single tap, where the effect occurs after the gesture. Continuous gestures can be shuffling dice while one hand mimicking shuffling action in the air. The effect occurs as the gesture happens.
- ***Complexity:*** Lastly, Complexity classifies gesture into Simple with one step only, and Compound with multiple steps, such as tapping private and then the public device to move a card out to shared space.

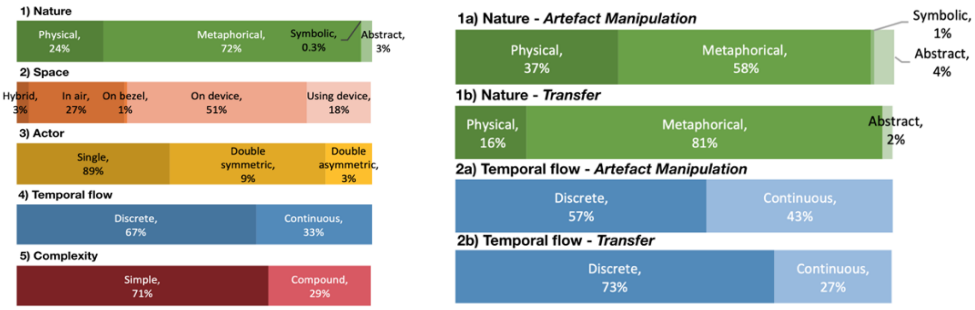


Fig. 3. (left) Overall classification breakdown of all gesture proposals; (right) distributions of Nature and Temporal flow dimensions under artefact manipulation and transfer referent (see Table 1).

4.2.2 *Classification Breakdowns.* The overall classification breakdown is shown in Figure 3 left. We observed more Simple gestures than Compound ones, as well as more Discrete gestures than Continuous ones. Similar to previous research studies [12,44,50], most gestures are Metaphorical in nature. Due to our study design, some referent-modality combinations are predetermined under the classification. For example, artefact manipulation referents such as shuffle cards are bound to be single actor only, and touch gestures must be performed on device only. For this reason, we took a finer look into the classification. We report some interesting comparisons found in our analysis below.

Comparing artefact manipulation (e.g., shuffle cards/dice) and transfer referents (e.g., move in/out a card, see Table 1) in terms of Nature (Figure 3 right), we observed a larger percentage of metaphorical gestures, but a smaller one for physical gestures in transfer, when compared to the artefact manipulation referents.

In Temporal flow, there is a larger percentage of continuous gestures in artefact manipulation than in transfer (Figure 4 right). In fact, 55% of mid-air gestures for artefact manipulation were continuous, also the highest proportion when compared to other referent category and modality combination (Table 3).

Table 3: Percentage of continuous gestures under all four referent category/modality combinations.

Modalities	Transfer Referent	Artefact Manipulation Referent
Touch	38%	30%
Mid-air	15%	55%

Comparing modalities for transfer referents under the Actor dimension (Figure 4), we observed that the elicited mid-air gestures involve a higher proportion of double actor than it is for touch gestures. As well, double asymmetric gestures were elicited in mid-air but not in touch gestures.

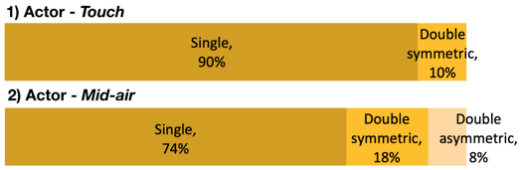


Fig. 4. Distribution of Actor for Transfer referent – Touch vs. Mid-air. More collaborative gestures (shown by a higher percentage in double symmetric and asymmetric) in Mid-air gestures than in Touch gestures.

4.3 Agreement Rate

We used AGATe 2.0 tool [46] for the modified agreement rate computation (Equation 1). The group sizes, i.e., the number of times a unique gesture was proposed, and the total no. of proposals for each referent-modality combination were used to compute the agreement rates. The full definition of the formula can be found in Vatavu and Wobbrock's work [46]. Agreement rates, confidence intervals, and significance are shown in Figure 5.

$$AR(r) = \frac{|P|}{|P|-1} \sum_{P_i \subseteq P} \left(\frac{|P_i|}{|P|} \right)^2 - \frac{1}{|P|-1} \qquad (1)$$

Overall agreement rate (AR) ranged from 0.061 (AR < 0.100, considered a low agreement [46]) to 0.545 (AR > 0.500, considered a very high agreement [46]), with a mean overall AR of 0.159. Informed by the low convergence of hand gestures from prior work [31], as expected, the average agreement for Touch gestures (ARtouch=0.215) was higher than that for Mid-air gestures (ARmid-air=0.101). This difference in agreement shows a wider variety of gesture proposals in the Mid-air modality than in the Touch modality. Participants also frequently expressed the difficulty to propose Touch gestures: "*touch is always so difficult, cos there's only a limited of things you can do*" (P20). Another contributing factor could the legacy bias from touch gestures with common mobile devices.

To understand the difference of agreement rates across modalities, we compared individual referents across modalities. The Touch gestures in five referents, namely *Give card, Exchange, Give chip, Flip,* and *Shuffle card*, have significantly higher agreement rates than their Mid-air counterparts. The significance is indicated by asterisks in Figure 5.

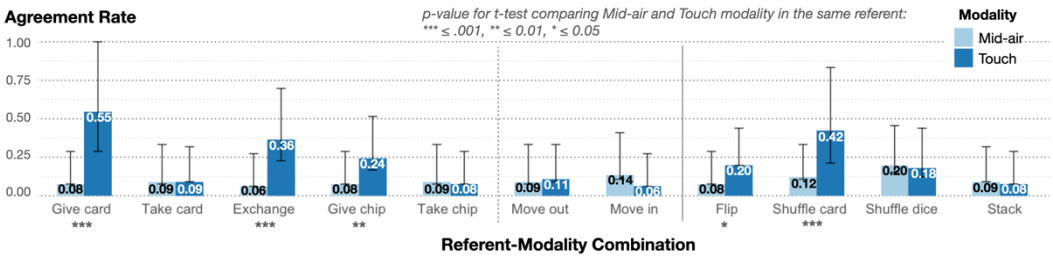
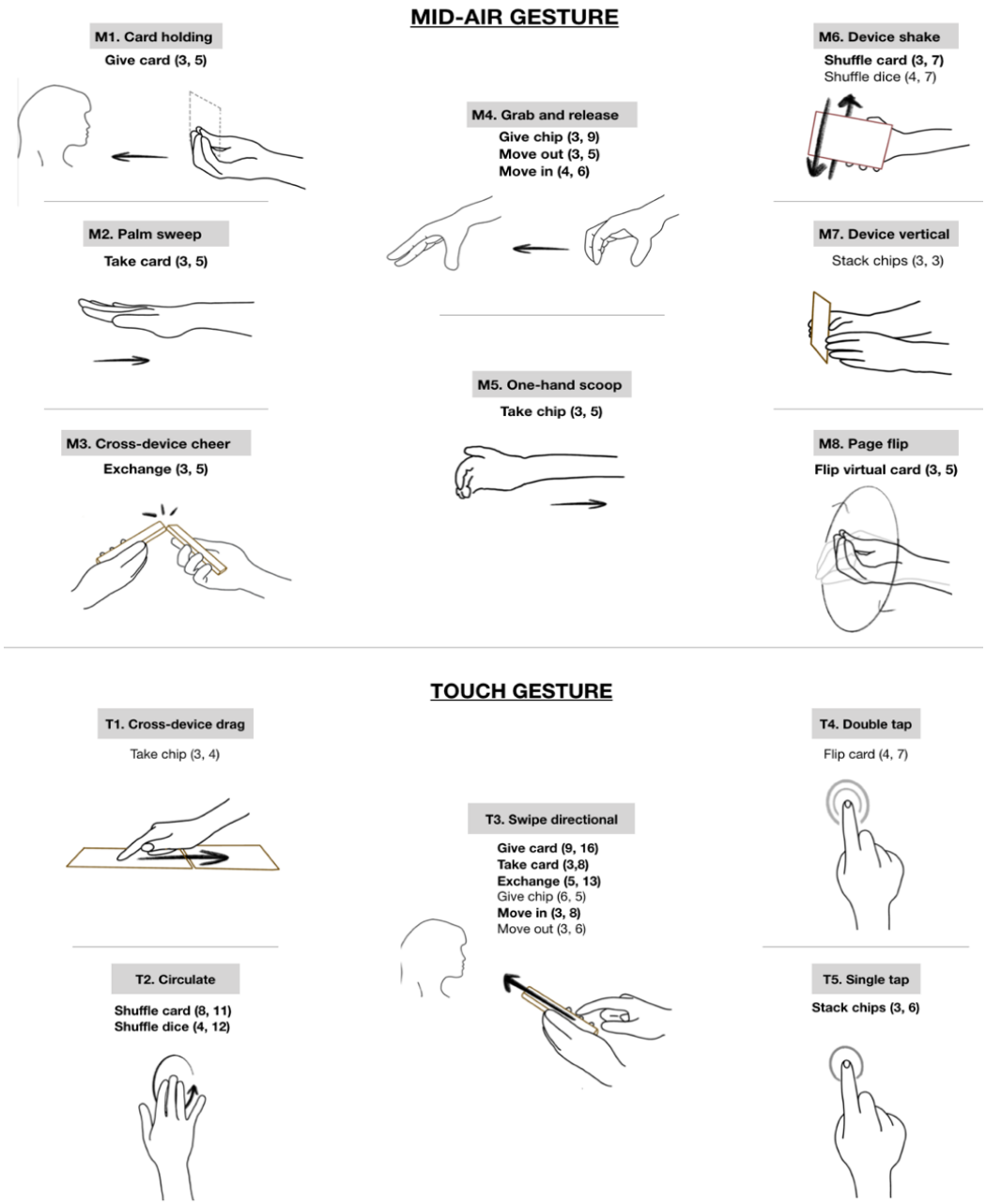


Fig. 5. Agreement rates and confidence intervals of the 22 referent-modality combinations.



- **Bold referent** means the consensus gesture is also the *most frequently proposed*.
- In the bracket, the first number is *no. of dyads chosen the gesture as favorite.* The second number is the *no. of times the gesture was proposed* for that referent-modality combination.
- Note that only one variation of the gesture proposal is presented here for simplicity.
- For arrow moving direction, the actor is on the right side. Graphics showing right hand as dominant hand.

Fig. 6. Consensus gesture set – 13 gestures for 22 referent-modality combinations.

Two dichotomous pairs – *Give card touch / Take card touch* and *Give chip touch / Take chip touch* show a significant difference in agreement rates. However, we did not see a specific trend in these differences. In addition, as the referents were presented in random order during the elicitation, some participants chose a different gesture instead of the opposite of the dichotomous gesture sometimes because they might have forgotten what they proposed before as they confessed. However, we observed some participants deliberately recalled and matched the gestures with the gestures proposed for opposite or similar action previously.

4.4 Consensus Gesture Set

After gesture classification and agreement rate computation, we developed the consensus gesture set (Figure 6) by first selecting the most favourite gestures in each referent-modality combination. When there was a tie, we chose the most frequently mentioned gesture as the consensus. We were able to resolve all ties with this procedure. It is worth noting that the above selection procedure could result in cases where the final consensus gesture was the most favourite one but not the most frequently mentioned. In our dataset, there were 6 out of 22 such cases, where the consensus gestures were the second most frequently proposed. The remaining 16 out of 22 final gestures were the most favourite ones, and at the same time, the most frequently mentioned. There are conflicts within the gesture set, i.e., the same gesture for multiple referents. Conflicting gestures are retained because participants preferred using variations of the same gesture for similar referents to help with memorability and learnability. Swipe directional is the favourite Touch gestures for 6 out of 7 Transfer referents, in which 2 pairs are dichotomous referents, namely, *Give card* and *Take card*, *Move in* and *Move out*. This result also shows the same Touch gesture is applied regardless of the types of device, i.e., public or private device, involved. However, there is a higher variety of consensus Mid-air gestures for Transfer referents – 5 different gestures out of 7, with *Grab and Release* repeating three times in *Give chip, Move out,* and *Move in*. Using the same gesture for similar operation can also be seen in artefact manipulation referents. For example, Shuffle cards and dice, have the same consensus touch and mid-air gesture. The complete gesture set can be found in the Appendix.

4.5 Subjective Ratings

Participants rated different areas of the gesture elicitation (Figure 7) in a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree). In terms of the elicitation experience, participants understood the study requirement (mean=6.50, SD=0.65), and agreed with the final gestures chosen favourite (mean=6.38, SD=0.75). In general, participants found it fun to propose gesture by themselves (mean=6.25, SD=0.77) and with a partner (mean=6.55, SD=0.81). However, 7 participants found it difficult to suggest gestures (mean=3.41, SD=1.73). Regarding the final gestures chosen favourite, participants agreed that the set was a good fit for its purpose (mean=6.46, SD=0.58), and to a lesser extent, learnable (mean=5.67, SD=1.40) and easy to perform (mean=5.92, SD=0.95). Response was mixed for whether the gestures were tiring to perform (mean=3.5, SD=1.80).

We also asked participants about the possibility to adopt the proposed gestures in other settings. Participants strongly agreed that they would use gestures in a game with other people. They agreed on gesture in games to be *"fun and playful"* (P11) with friends but not with strangers. They generally agreed that they would use gesture in other casual settings, but attitudes vary widely for settings with other people, which is likely because proposed gestures were *"specific to a card game"* (P6). Participants welcomed using gestures in settings other than games when they

are discreet, can streamline workflow, or serve a clear purpose, such as for *"sharing something between two screens at work"* (P5). They were concerned about using gestures because they were not confident about the accuracy of sensing technologies or due to *"social awkwardness"*.

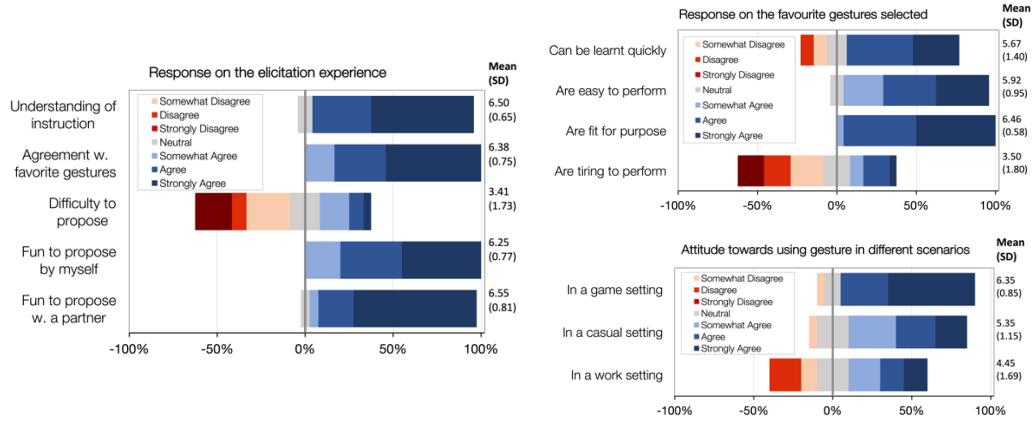


Fig. 7. Results of Post-study Questionnaire on the elicitation experience, favourite gestures selected and attitude towards using gestures in 3 scenarios. Likert-scale from 1 (strongly disagree) to 7 (strongly agree).

5 INTERACTION PATTERNS

In this section, we describe the interaction themes from the thematic analysis [6] of think-aloud elicitation sessions and exit interviews, complementing our quantitative elicitation results.

5.1 Relevance and Appropriateness of Gesture Input

All participants welcomed the idea of using gestures to control mobile devices in a board or card game, but to various extents. Most participants said they would use gestural input when the gestures are *"relevant"* and *"appropriate"* and are performed with people they already know: *"It would be great if I can do some hand gesture with my friends when we are playing the game"* (P20). Many were hesitant towards gesture usage beyond their social circle: *"I would be comfortable doing gestures with friends but not with strangers."* (P19). The most enthusiastic participants viewed gesture usage as *"safe to use with anyone"* as it does not *"require any level of intimacy"* (P10), and were *"confident that I can use it [gesture] in every situation"* (P16).

Participants recognised the potential of gestures to streamline the game flow, especially in the Transfer referents. As participants were highly reluctant towards *"a [menu] panel full of buttons"* (P22) and *"tapping [on menu]"* (P23). Participants believed gestures should *"assist but not obstruct the game"* (P15), and preferred gestures such as *"just drag it to their face"* to give a card to a partner because it is *"easier"* and more *"convenient"* than *"clicking lots of buttons"* (P24). A participant went further to question the use of menu bar, "if we use it [a menu bar] all the time, what's the point of having this [mobile device]?" (P7).

However, participants were also concerned about using gestures in real life. Some participants expressed concern over gesture recognition. P24 viewed that hand gestures in air *"could mean anything, [are] harder to recognise, [and] easier to get it wrong"*. P20 was "hesitant to use gesture" *"if there is no feedback, no confirmation"* of the gestures performed. Participant age might also affect how they viewed gestural control, as P18 expressed, *"touch is what I'm used to – being an old person"*.

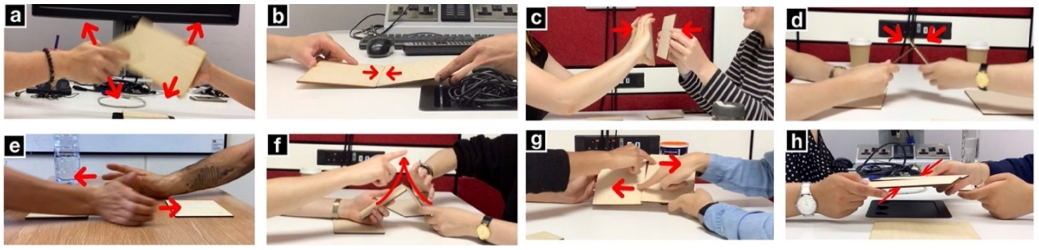


Fig. 8. Examples of collaborative gestures. Arrows indicate moving direction. (a) shake devices in air with front of device facing each other; (b) connect devices flat on table; (c) connect devices back-to-back in air; (d) cheer with devices; (e) one hand push on the side; (f) finger swipe on connected devices, finger ends in air towards receiver; (g) connect devices flat in air, finger drag across devices; (h) parallel devices in contact sideway.

5.2 Collaborative Gestures Fostering Interaction Between Players

Participants preferred gestures that gave them a sense of achievement and foster interaction between players. One of the contributing factors to a gesture being selected as favourite was the degree of interaction that the gesture offers: P22 asked, *"Do you like more interaction, or do you want to stay on your own?"* P21 and P22 eventually agreed on a collaborative gesture – dragging across two devices, instead of swiping on their own devices, for exchanging cards (Figure 8g). P14 commented collaborative and interactive gestures allow *"build[ing] more connection, [and being] more integrated with your team"*, which was considered as an important element of multiplayer games. This rationale was behind consensus collaborative gestures such as cheering using two devices (Figure 8d) and touching with two parallel devices (Figure 8h).

Although some Touch gestures include collaboration, participants preferred Mid-air over Touch potentially because of the more prominent need to cooperate for a gesture in air than it is on screen. As P15 commented, *"[mid-air gesture] can bring a more interactive experience rather than purely touch-based"*. P10 summarised this preference best: *"[Mid-air gesture] is more interactive, feel more physical. The purpose of the game is to have fun, and maybe bumping two devices is more fun than swiping because it's a collaborative game, you want to have another thing to give to your friend. If we use touch, it feels like we are sitting together but we are not interacting, but with gestures like bumping, we can interact."* Other examples of collaborative gesture proposals are shown in Figure 8. quire any level of intimacy" (P10) and were *"confident that I can use it [gesture] in every situation"* (P16).

5.3 Benefits of Situation Awareness

Participants recognised realistic gestures can assist the game by forming shared situation awareness [14] in all players. All participants preferred realistic gestures and often translated gestures used in real board and card games directly into gestural input for manipulating virtual objects. P7 explained her mental model: *"This is what I would do [gesturing picking a card from an accordion held by P8], which is how I would do it here [gesturing the pinching action on P8's device]"* (Figure 9a). Another example is double tapping as a metaphor to pressing buttons on a dice roller. Participants considered being *easy to understand* and *universal* as important features of gestures. As P20 explained: *"shuffling a deck of card in air would be instantly recognized, therefore effective"* (Figure 9b).

Another important aspect of situational awareness is that it allows learning and teaching opportunities through gestural input. In the free-form response from the post-study questionnaire, participants expressed that the gestures representative of gameplay in real life were *"especially useful when the game is played in a big setting with many people"* (P15) and *"would enjoy teaching individuals [the gestures chosen]"* (P11). However, there were also concerns about the similarity between the gestures proposed that would be confusing or ambiguous to both players (especially novice) and sensors.

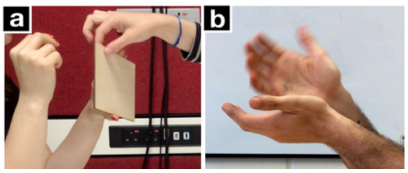


Fig. 9. Examples of realistic gestures: (a) pinch a card from an imaginary accordion on the partner's device; (b) bare hands shuffling imaginary cards in air.

5.4 Gestures Providing Fun, Excitement, and Suspense to the Game

Participants were enthusiastic about proposing artefact-manipulative gestures and frequently described them as *fun*. Mid-air gestures such as shuffling virtual dice with bare hands in air were *"part of the fun of the game"* (P5) and were *"real and fun"* (P9). Tapping palm repeatedly with the side of the device to shuffle cards *"actually feels like I'm shuffling"* (P13). Similarly, the same enthusiasm can be seen in repetitive Touch gesture: *"you want to tap really fast [on screen] to get all the chips"* (P11). In addition, participants merited the ease of touch to manipulate artefacts when compared to: *"it [fingers circulating on screen] still gets the fun, more effort [required] to put down the device and shuffle [with hands in air]"* (P13).

Repeated and exaggerating gestures offer a sense of satisfaction and at times, tension to the game. Hugging imaginary chips on the *side "makes me feel more like winning"* (P9). P13 described it was *"quite fun to put device on top of other to take chips"* because the gesture allowed him to interact with his partner, teasing *"I take your chips now!"*.

Other gesture such as swiping in air along the digital cards to choose a card from the partner also *"adds excitement and suspense"* (P12). Participants viewed this tension harder to achieve with touch-only interaction. However, this fun part of the game might remain to artefact manipulation but do not apply to other functional game tasks. P8 explained in detail: *"because it is fun to shuffle then I want to interact with it [the device], but if it is other action like flipping a card, I just want to see it [the card], then I want it easy"*.

5.5 Social Etiquette

We observed the tendency to avoid reaching close to the region of partner's device in the one-way transfer of objects to or from the partner. This pattern was seen in both Touch and Mid-air gestures. For taking chips from the partner, participants proposed "scooping with one hand", either on the side, or at centre ending in front of a partner's device without touching (Figure 10a, 10b). One of the reasons was that participant preferred the convenience when the hand does not need to reach too far: *"I don't want to touch the other device, and then scoop it all the way, like if I need to get from many players"* (P9). Some participants viewed hand going too far as impolite or invasive and they *"don't want to be mean about it"* (P5), as *"it might be impolite or invading other people's zone"* (P7). Using only one hand was preferred over both hands because it was *"easier"* and that it might be *"awkward to have two hands if you are reaching towards other people"* (P5).

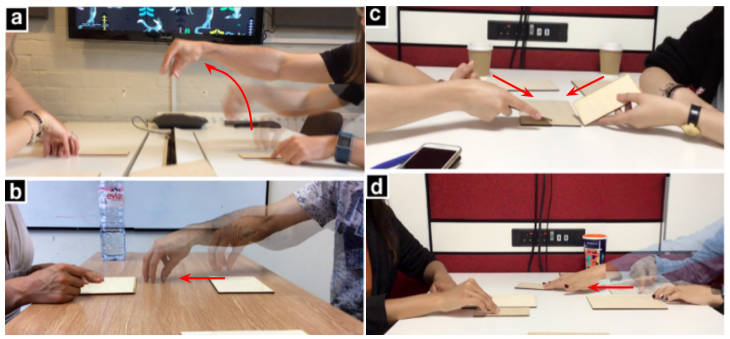


Fig. 10. Gestures for transferring game artefacts. Note that in the first two gestures participants avoid reaching to partners' devices. (a) scooping action ends in air in front of the partner; (b) scooping on bezel ending in front of the partner's device; (c) tilt private device to touch public device; (d) use of a secondary device to transfer object from private device to partner.

However, this avoidance did not occur in gestures for transferring digital objects between private and public device. Participants were willing to reach for or touch the public device. In the Move out referent, for example, a Touch gesture proposed was tapping private and then the public device, and a Mid-air gesture proposed was tilting and tapping public device with the private device – both involve contacting the shared device (Figure 10c). One dyad went further to propose using a secondary device as a medium to transfer object between players instead of directly transferring (Figure 10d). In short, participants considered social etiquette when using gestures for a referent that involves a partner.

5.6 Impact of Legacy Bias

Existing common touch-based interaction techniques in mobile phones, laptops or desktops were repeatedly referred to throughout the study. Tap and hold, then drag was selected as the favourite touch gesture for stacking digital chips because participant viewed it as a common gesture on phone – *"it's like we tap and hold to choose the app, then drag it to another place"* (P8). P7 referred to *"a new function on Mac Keynote"* when proposing gesture for Stack chips because the gesture "can stack things by dragging one on another".

It was also seen that existing touch-based interaction has formed a user habit in interpreting the outcome of a gesture. Participant preferred double tap over a mid-air gesture for flipping a card because of the functional difference in these two actions in computer operating systems – *"one tap is usually for selection; double tap is for opening something"* (P19). The expected outcome of tapping is already imprinted in the users' minds. As such, they preferred consistency between tabletop games and productivity work in a computer. The final verdict was that *"double tap was imprinted in my already familiarity with Mac"* (P20).

Although we did not employ any graphical user interface nor show any traditional game platforms before the elicitation, menu-based panel control was repeatedly proposed in both Touch and Mid-air gesture. In cases where it was a struggle to think of new gestures, participants often resorted to menu-based interaction to fulfil the minimum requirement. However, participants often showed disappointment when they do so*: "but it's boring"* (P24). Participants actively stayed away from menu-based interaction as they opposed the idea of using "a panel of buttons to control" because *"you don't really want so many buttons here."* (P22). Another participant questioned: *"we can do this [menu command] on a computer as well, what's the point of using this [referring to the plywood device]?"* (P8).

6 DISCUSSION

Guided by the research gap in user-defined interaction techniques for co-located mobile games, our work explored attitudes towards gestural control and generated a consensus gesture set through our gesture elicitation. Combining the high enjoyment ranking for the gesture proposal experience, the high preference ranking in using gesture in mobile games, and the feedback from the exit interviews, we observed that our participants were welcoming towards gestural control for co-located mobile games. In the following, we discuss the implication for design and relation to earlier work.

On Input Modalities and Use in Interaction Design. While our work does not attempt to provide a definitive guide on when to employ touch or mid-air gesture, our agreement rates and qualitative findings give insights into the suitable modality (touch, mid-air) for a referent, which can be a reference point for interaction designers when designing gestures for different game tasks. For example, when designing for the transfer of game object between players, leveraging on our identified interaction patterns might result in choosing gestures that require involvement of both players to enhance physical interaction. As mentioned before, we observed significantly higher agreement rates in Touch compared to Mid-air gesture in some of the Transfer referents, which hinted that Touch gesture might be suitable for referents that involve transferring objects. This might be because touch allows accurate selection of digital objects and receiver, which is harder to achieve in mid-air gesture. Similarly, we observed comparable agreement rates in Artefact Manipulation referents across modalities. Combining with the qualitative findings that participants regard artefact manipulation as *"part of the fun"* when they proposed realistic gestures such as shuffling dice with bare hands in air, gesture designer might consider using mid-air or motion gestures that might enhance game engagement. However, this implication should be considered with care and examined for individual cases, because agreement rate alone has its own pitfalls [38], therefore does not provide a definitive guide on gesture design.

Design to Assist and Amuse. From our exit interviews and think-aloud, we saw that participants were welcoming when the gesture was assistive to the game. A similar finding was also observed in prior work [27] where user preferred cooperative gestures that are necessary to aid the tasks. The use of mid-air or motion gestures allows all players and spectators to recognise game events concurrently more easily as they happen, which is hard to achieve with solely touch-based interaction on screen. This is possibly due to the implicit nature of touch-based input, in which the situation awareness relies more on the game content itself or by verbal communication among co-located players. Whereas mid-air and motion gestures are more noticeable to players and spectators compared to the game content on mobile device screens. As screen content might be less accessible or comprehensible to novice players and spectators, being able to observe game events from hand gestures allows all stakeholders to engage with the game. As players rely on communicative cues to understand the current state in the game [48], situation awareness is especially important in a co-located game setting. We, therefore, recommend designers to consider using gestures to enhance situation awareness. Another route to design gestures was when the gestures add enjoyment to the game. Therefore, gestures could be entertaining inherently, or designed to build interaction between players. To sum up, game designers might take advantage of modalities beyond touch to enrich the co-located gaming experience for all stakeholders, as well as to be inclusive to novice players with the educational opportunities provided by gestural input.

Consistent Gestures Across Devices. When designing for mobile game interaction, it is important to design a cohesive gesture set that is available across different form factors. In our elicitation study, participants often proposed a gesture suitable for the device formfactor they were holding (described in Apparatus, there were plywood boards of three sizes), but then pinpointed the limitation the gesture would have on other mobile devices that they would propose another gesture or modify the existing gesture in order to suit mobile devices of different form factors. For example, palm tapping on screen was viewed as inaccurate and inconvenient to perform on small screens, leading participants to re-suggest for the gesture to use only three fingers. Another example is the motion gesture of shaking the device vertically, for which participants thought it might be difficult to perform with larger devices due to heavy weight. In light of this, a strand of future work can investigate co-located mobile game gestures that are consistent across mobile device type. Similar research topics started to gain momentum in other application areas like Dingler et al.'s work [12] in gesture elicitation for active reading.

Explicit and Implicit Feedback. From our classification breakdowns and qualitative findings, we observed the need for continuous feedback, especially in artefact manipulating referents. As mentioned earlier, mid-air gestures for artefact manipulation have the highest percentage of continuous gesture proposals when compared to other referent categories and modality combinations. Similarly, participants showed hesitance to mid-air gestures because of the lack of feedback: *"[unsure about mid-air gesture] if I don't see what happen to it, but for touch, I know what I'm touching cos I can see it and feel it"* (P20). While touch gestures ensure direct contact with the digital object through a capacitive screen, purely performing hand gestures in air might confuse users because of the lack of implicit feedback. We recommend designers to explore explicit visual feedback to guide users in their action, as feedback can help users build their understanding of the gesture-based system. As in mid-air gestures for artefact manipulation referents such as hands shuffling imaginary card in air, designers might consider using the device as a form of feedback. An example would be one of the gesture proposals to shuffle cards by tapping the side of the device on palm repeatedly, which gives tactile feedback.

Limitations and Future Work. Although our aim is to explore gestural control for co-located multiplayer mobile games, our study only looked into interaction with dyads of participants. In real life, there will often be more than two players in board or card games. Participants also expressed worries on the accuracy of mid-air pointing from a distance when the number of players around the table is too large. A larger number of players might impact both situation awareness and recognition accuracy by technology. Some were also concerned about the difficulty in direct touch interaction if there were many players. In addition, our elicitation was conducted in a formal setting which is different from real-life leisure gameplay. Therefore, future work might yield different insights through user elicitation with more participants in the wild. For future studies, it would also be interesting to further investigate of how gender might have impact on gestures can be a potential line of inquiries, especially for game tasks that require physical interaction between players. Furthermore, as this elicitation study was conducted with participants sitting opposite of each other, it would be interesting to find out if and how varying seating arrangements (such as sitting adjacent to one another) might have an impact on the kinds of gestures. Another next step for this study could be to implement consensus gestures on mobile devices for game testing with groups of participants. The results could add to the iterative design of gesture and validation of the consensus set.

7 CONCLUSIONS

To inform the use of gestures in co-located mobile games, we explored user-defined gestures for co-located mobile gaming with twelve dyads through a gesture elicitation study. We presented a consensus as well as a touch and mid-air gesture set for co-located mobile game tasks. Our agreement rate analysis revealed a higher consensus in touch gesture proposal than mid-air. We also classified the gestures according to five dimensions, showing the need for continuous feedback when designing gestures. Our thematic analysis of interviews and think-aloud sessions showed a positive response to touch and mid-air gesture control for games, which indicates the potential of gesture control in mobile gaming. Collaborative gestures and artefact manipulation gestures have the potential to contribute making co-located mobile games more engaging. Based on our findings, we discussed guidelines for game gesture design and implications for sensing technologies. Overall, our work serves as a starting point for game gesture designers, as well as a reference for future research in this field.

APPENDICES

A.1 Complete Unique Gesture Set

Referent	Touch	Mid-air
Give card	**swipe directional (9, 16),** double tap, *drag directional, drag menu,* finger circulate, finger walk across devices, tap on directional, *tap menu*	**card giving (3, 5),** back hand sweep, CD cheer, CD connect, CD device tap, device flick, finger drag, finger flick, grab and release, palm swipe, pinch out, secondary device, *swipe*
Take card	**swipe directional (3, 8),** CD connect, *CD drag,* CD swipe, CD tap, *drag menu,* pinch on screen, pinch out, single tap, *swipe directional,* swipe fan, *tap menu,* drag directional	**palm sweep (3, 5),** card giving, CD cheer, CD connect, CD device tap, CD tap, CD tap, device flick, device point, finger rub and hand swipe, grab and release, index swipe, *palm sweep, pinch out,* secondary device, swipe and perpendicular, swipe fan
Exchange	**swipe directional (5, 13),** *CD drag,* circulate, double tap, drag menu, fingerprint, peace sign cross, *swipe directional,* tap and hold, *tap menu*	**CD cheer (3, 5),** card giving, *CD cheer, CD connect, CD device flick,* CD device shake, device flick, device point, device tilt, finger flick, finger twirl ,grab and release, high five, page flip, palm slide, palm sweep, peace sign cross, pinch out, screen down, secondary device, swipe directional
Give chip	swipe directional (6, 5), all finger centralised drag, all finger centralised tap, CD swipe, chip giving, double tap, ***drag directional,*** drag menu, finger drumming, repeated tap, side of palm swipe, directional, *swipe directional, tap menu*	**grab and release (3, 9),** *1 hand hug,* 2 hand gather, *2 hand hug,* CD connect, CD device tap, device air tap, device tilt, *grab and release,* index point, palm sweep, palm swipe, secondary device, swipe directional
Take chip	CD drag (3, 4), all fingers pinch in, *CD drag,* CD swipe, double tap, ***drag directional,*** drag menu, pinch on screen, press and hold, repeated tap, *swipe directional,* tap menu	**1 hand hug (3, 5),** ***1 hand hug,*** 2 hand hug, CD cheer, *CD connect,* CD shake??, *CD swipe, grab and release,* palm sweep, palm swipe, perpendicular, swipe directional
Move out	**swipe directional (3, 8),** CD drag, CD swipe directional, *CD tap,* device tilt, double tap, *drag directional,* drag menu, ***swipe directional,*** swipe sideway, tap directional, tap menu	**grab and release (3, 5),** 2 hand card dealing, CD all fingers centralised tap, *CD connect,* CD device shake, *CD device tap, CD flip,* device shake, device tilt, finger flick, finger point, ***grab and release,*** *palm swipe,* pinch in pinch out, secondary device

Move in	swipe directional (3, 5), CD connect drag, *CD drag,* CD swipe, ***CD tap,*** double tap, drag, drag directional, drag menu, finger walk, fingerprint tap, pinch in pinch out, single tap, *swipe directional,* swipe sideway, tap menu, zigzag	**grab and release (4, 6),** card holding, CD all fingers centralised tap, *CD connect,* CD device flick, CD device flip, *CD device tap,* CD device tilt, CD drag, CD tap, CD wave, delete, device tap, finger point, *grab and release,* palm sweep, palm swipe, pinch out, swipe directional
Flip	double tap (4, 7), *double tap,* drumming, finger twist, *half circle, single tap,* ***swipe,*** tap and hold	**page flip (3, 5),** clap, ***device flick,*** *device flip,* device tilt, finger flip, finger rub, fist bump, flick device, flip device, hand top and bottom, *index twirl,* ***page flip,*** *palm flip,* palm sweep
Shuffle card	***circulate (8, 11),*** *circulate,* device flip, diagonal drag, double tap, finger drumming, palm tap, *repeated swipe, single tap,* swipe, swipe and tap, swipe sideway, tap and hold	**device shake (3, 7),** 1 hand shuffle card, *2 hand shuffle card,* all finger centralised lift, card breaking, device as card shuffle, device flick, device flip, *device shake,* finger snap, finger twirl, grab and release, hand wash, palm double tap on side, repeated pinch out, *wobble*
Shuffle dice	***circulate (4, 12),*** *circulate, double tap,* finger drumming, finger twist, flick, repeated swipe, *repeated tap, single tap,* swipe up and down, tap and hold	device shake (4, 7), ***1 hand dice,*** *2 hand dice,* all fingers centralised drag, all fingers centralised lift, circulate, *device shake,* device up and down, finger twirl, grab and release, hand shake, wobble
Stack	**single tap (3, 6),** 2 hand pinch in, all fingers centralised, check mark, circulate, *drag chip,* palm tap, *pinch in, repeated tap, single tap,* square *swipe up* tap and hold	device perpendicular (3, 3), 1 bent finger lift, 1 hand hug, *2 hand hug,* 2-finger c shape air tap, all finger centralised tap, ***all fingers centralised lift,*** all fingers centralised tap, device lift, *device perpendicular,* device shake, finger tap, finger twirl, grab, palm stack, palm swoop, palm tap, repeated palm swipe, tap on side of device
Bold items are the gesture most frequently proposed in the referent-modality combination. First items are the most favourite gestures. In the bracket, the first number is the number of participants chosen the gesture as favourite, the second number is the no. of times the gesture was proposed.		

A.2 Consensus-distinct Ratio and Max-consensus Ratio

We calculated consensus-distinct ratio and max-consensus ratio [25] to explore any patterns form the data. We prioritised the findings presented in the main paper over these metrics. Note that 't' is the distinct threshold.

	consensus-distinct (t=3) (touch)	*consensus-distinct (t=3) (mid-air)*	*max-consensus (touch)*	*max-consensus (mid-air)*
Give card	0.867	0.308	0.750	0.250
Take card	0.733	0.310	0.250	0.250
Exchange	0.806	0.438	0.583	0.250
Give chip	0.581	0.536	0.417	0.250
Take chip	0.593	0.655	0.250	0.250
Move out	0.633	0.643	0.250	0.250
Move in	0.516	0.387	0.250	0.333
Flip	0.867	0.677	0.333	0.250
Shuffle card	0.645	0.600	0.667	0.250
Shuffle dice	0.774	0.677	0.333	0.333
Stack	0.758	0.375	0.250	0.250

REFERENCES

[1] Roland Aigner, Daniel Wigdor, Hrvoje Benko, Michael Haller, David Lindbauer, Alexandra Ion, Shengdong Zhao, and Jeffrey Tzu Kwan Valino Koh. 2012. Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI. *MSR-TR-2012-111 Technical Report.* Retrieved May 23, 2022 from https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-types-for-hci/

[2] Alexandre Almeida and Ana Alves. 2017. Activity recognition for movement-based interaction in mobile games. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '17), 1–8. https://doi.org/10.1145/3098279.3125443

[3] Saskia Bakker, Debby Vorstenbosch, Elise van den Hoven, Gerard Hollemans, and Tom Bergman. 2007. Weathergods: Tangible Interaction in a Digital Tabletop Game. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (TEI '07), 151–152. https://doi.org/10.1145/1226969.1227000

[4] Rob van Bekkum, Thijs L.M. Brands, Soheil S. Jahanshahi, Aidan C.A. Mauricio, Joost J.E. Oorschot van, Fanny Lie, Ben Kybartas, and Rafael Bidarra. 2014. Taxi trouble: communication is key. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play* (CHI PLAY '14), 387–390. https://doi.org/10.1145/2658537.2662994

[5] Po Bhattacharyya, Radha Nath, Yein Jo, Ketki Jadhav, and Jessica Hammer. 2019. Brick: Toward A Model for Designing Synchronous Colocated Augmented Reality Games. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI '19), 323:1-323:9. https://doi.org/10.1145/3290605.3300553

[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2: 77–101. https://doi.org/10.1191/1478088706qp063oa

[7] Philipp Brauner, Julia van Heek, Martina Ziefle, Nur Al-huda Hamdan, and Jan Borchers. 2017. Interactive FUrniTURE: Evaluation of Smart Interactive Textile Interfaces for Home Environments. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (ISS '17), 151–160. https://doi.org/10.1145/3132272.3134128

[8] Liselott Brunnberg. 2004. The Road Rager: making use of traffic encounters in a mobile multiplayer game. In *Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia* (MUM '04), 33–39. https://doi.org/10.1145/1052380.1052386

[9] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer. 2016. User Elicitation on Single-hand Microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 3403–3414. https://doi.org/10.1145/2858036.2858589

[10] Chi Tai Dang and Elisabeth André. 2010. Surface-poker: Multimodality in Tabletop Games. In *Extended Abstracts of ACM International Conference on Interactive Tabletops and Surfaces* (ITS '10), 251–252. https://doi.org/10.1145/1936652.1936701

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

[12] Tilman Dingler, Rufat Rzayev, Alireza Sahami Shirazi, and Niels Henze. 2018. Designing Consistent Gestures Across Device Types: Eliciting RSVP Controls for Phone, Watch, and Glasses. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), 419:1-419:12. https://doi.org/10.1145/3173574.3173993

[13] Guiying Du, Auriol Degbelo, Christian Kray, and M. Painho. 2018. Gestural interaction with 3D objects shown on public displays: an elicitation study. *Interaction Design and Architecture(s)* 38: 184–202. https://doi.org/10.55612/s-5002-038-010

[14] Megan Ellis. 2017. 5 Mobile Games You Control in Crazy Ways. Retrieved May 23, 2022 - from https://www.makeuseof.com/tag/mobile-games-crazy-control/.

[15] Leila Homaeian, Nippun Goyal, James R. Wallace, and Stacey D. Scott. 2018. Group vs Individual: Impact of TOUCH and TILT Cross-Device Interactions on Mixed-Focus Collaboration. In *Proceedings*

of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), 73:1-73:13. https://doi.org/10.1145/3173574.3173647

[16] itch.io. 2022. Top games for Android with Accelerometer support. itch.io. Retrieved May 23, 2022 - https://itch.io/games/input-accelerometer/platform-android.

[17] Ahmed S. Khalaf, Sultan A. Alharthi, Igor Dolgov, and Z O. Toups. 2019. A Comparative Study of Hand Gesture Recognition Devices in the Context of Game Design. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (ISS '19), 397–402. https://doi.org/10.1145/3343055.3360758

[18] Seungki Kim, Donghyeon Ko, and Woohun Lee. 2017. Utilizing Smartphones as a Multi-Device Single Display Groupware to Design Collaborative Games. In *Proceedings of the 2017 Conference on Designing Interactive Systems - DIS '17*, 1341–1352. https://doi.org/10.1145/3064663.3064716

[19] Christian Kray, Daniel Nesbitt, John Dawson, and Michael Rohs. 2010. User-defined Gestures for Connecting Mobile Phones, Public Displays, and Tabletops. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (MobileHCI '10), 239–248. https://doi.org/10.1145/1851600.1851640

[20] Janne Lautamäki and Riku Suomela. 2008. Using player proximity in mobile multiplayer games: experiences from Sandman. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology* (ACE '08), 248–251. https://doi.org/10.1145/1501750.1501807

[21] Oleksandr Lobunets and Wolfgang Prinz. 2011. Evaluating a smart working environment with a digital card game prototype. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (CSCW '11), 673–676. https://doi.org/10.1145/1958824.1958942

[22] Tony Manninen. 2004. Interaction Forms and Communicative Actions in Multiplayer Games. Retrieved May 23, 2022 - from http://www.gamestudies.org/0301/manninen/.

[23] Keenan R. May, Thomas M. Gable, and Bruce N. Walker. 2017. Designing an In-Vehicle Air Gesture Set Using Elicitation Methods. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (AutomotiveUI '17), 74–83. https://doi.org/10.1145/3122986.3123015

[24] Erin McAweeney, Haihua Zhang, and Michael Nebeling. 2018. User-Driven Design Principles for Gesture Representations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), 1–13. https://doi.org/10.1145/3173574.3174121

[25] 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* (ITS '12), 95–104. https://doi.org/10.1145/2396636.2396651

[26] Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, c. schraefel, and Jacob O. Wobbrock. 2014. Reducing legacy bias in gesture elicitation studies. *interactions* 21, 3: 40–45. https://doi.org/10.1145/2591689

[27] 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* (CHI '06), 1201–1210. https://doi.org/10.1145/1124772.1124952

[28] Michael Nielsen, Moritz Stoerring, Thomas Moeslund, and Erik Granum. 2003. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. 409–420. https://doi.org/10.1007/978-3-540-24598-8_38

[29] Steve Orzel. 2015. In the Lab: Designing Myo. Steve Orzel discusses the industrial design of the Myo armband. Retrieved May 23, 2022 - from https://developerblog.myo.com/myo-industrial-design/.

[30] Thomas Plank, Hans-Christian Jetter, Roman Rädle, Clemens N. Klokmose, Thomas Luger, and Harald Reiterer. 2017. Is Two Enough?! Studying Benefits, Barriers, and Biases of Multi-Tablet Use for Collaborative Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17), 4548–4560. https://doi.org/10.1145/3025453.3025537

[31] Roman Rädle, Hans-Christian Jetter, Mario Schreiner, Zhihao Lu, Harald Reiterer, and Yvonne Rogers. 2015. Spatially-aware or Spatially-agnostic? Elicitation and Evaluation of User-Defined Cross-Device

Interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 3913–3922. https://doi.org/10.1145/2702123.2702287

[32] Isabel Benavente Rodriguez and Nicolai Marquardt. 2017. Gesture Elicitation Study on How to Opt-in & Opt-out from Interactions with Public Displays. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (ISS '17), 32–41. https://doi.org/10.1145/3132272.3134118

[33] Melissa J. Rogerson, Martin Gibbs, and Wally Smith. 2016. "I Love All the Bits": The Materiality of Boardgames. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 3956–3969. https://doi.org/10.1145/2858036.2858433

[34] Melissa j. Rogerson, Lucy A. Sparrow, and Martin R. Gibbs. 2021. Unpacking "Boardgames with Apps": The Hybrid Digital Boardgame Model. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (CHI '21), 1–17. https://doi.org/10.1145/3411764.3445077

[35] 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* (CHI '11), 197–206. https://doi.org/10.1145/1978942.1978971

[36] Douglas Schuler and Aki Namioka (eds.). 1993. *Participatory Design: Principles and Practices*. CRC Press, Hillsdale, N.J.

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

[38] Alireza Sahami Shirazi, Tanja Döring, Pouyan Parvahan, Bernd Ahrens, and Albrecht Schmidt. 2009. Poker Surface: Combining a Multi-touch Table and Mobile Phones in Interactive Card Games. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '09), 73:1-73:2. https://doi.org/10.1145/1613858.1613945

[39] Chaklam Silpasuwanchai and Xiangshi Ren. 2015. Designing concurrent full-body gestures for intense gameplay. *International Journal of Human-Computer Studies* 80: 1–13. https://doi.org/10.1016/j.ijhcs.2015.02.010

[40] Mikael B. Skov, Jesper Kjeldskov, Jeni Paay, Heidi P. Jensen, and Marius P. Olsen. 2015. Investigating Cross-Device Interaction Techniques: A Case of Card Playing on Handhelds and Tablets. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* (OzCHI '15), 446–454. https://doi.org/10.1145/2838739.2838763

[41] Dorothé Smit, Bernhard Maurer, Martin Murer, Jens Reinhardt, and Katrin Wolf. 2019. Be the Meeple: New Perspectives on Traditional Board Games. In *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction* (TEI '19), 695–698. https://doi.org/10.1145/3294109.3295657

[42] Yanke Tan, Sang Ho Yoon, and Karthik Ramani. 2017. BikeGesture: User Elicitation and Performance of Micro Hand Gesture as Input for Cycling. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '17), 2147–2154. https://doi.org/10.1145/3027063.3053075

[43] Shan-Yuan Teng, Mu-Hsuan Chen, and Yung-Ta Lin. 2017. Way Out: A Multi-Layer Panorama Mobile Game Using Around-Body Interactions. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '17), 230–233. https://doi.org/10.1145/3027063.3048410

[44] 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 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 3327–3336. https://doi.org/10.1145/2702123.2702214

[45] 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* (CHI '14), 4107–4116. https://doi.org/10.1145/2556288.2557373

[46] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM*

Conference on Human Factors in Computing Systems - CHI '15, 1325–1334. https://doi.org/10.1145/2702123.2702223

[47] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. 2018. Gesture Elicitation Studies for Mid-Air Interaction: A Review. *Multimodal Technologies and Interaction* 2, 4: 65. https://doi.org/10.3390/mti2040065

[48] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. 2019. Frame-Based Elicitation of Mid-Air Gestures for a Smart Home Device Ecosystem. *Informatics* 6, 2: 23. https://doi.org/10.3390/informatics6020023

[49] Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the guessability of symbolic input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '05), 1869–1872. https://doi.org/10.1145/1056808.1057043

[50] 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* (CHI '09), 1083–1092. https://doi.org/10.1145/1518701.1518866

[51] Taoshuai Zhang, Jie Liu, and Yuanchun Shi. 2012. Enhancing Collaboration in Tabletop Board Game. In *Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction* (APCHI '12), 7–10. https://doi.org/10.1145/2350046.2350050

Received February 2022; revised May 2022; accepted June 2022.