Phone Sleight of Hand: Finger-Based Dexterous Gestures for Physical Interaction with Mobile Phones

Yen-Ting Yeh
y6yeh@uwaterloo.ca
Cheriton School of Computer Science,
University of Waterloo
Canada

Fabrice Matulic
fmatulic@preferred.jp
Preferred Networks Inc.
Japan

Daniel Vogel
dvogel@uwaterloo.ca
Cheriton School of Computer Science,
University of Waterloo
Canada

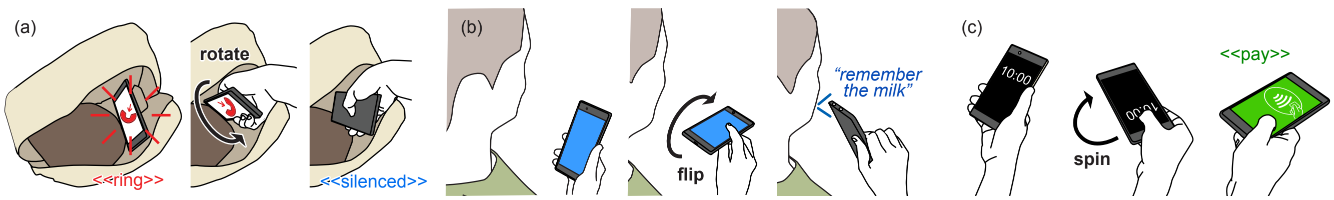

Figure 1: Examples of dexterous phone gestures: (a) half rotate for eyes-free actions, like silencing an incoming call; (b) half flip to activate a voice app, like note dictation; (c) half spin to open a dedicated app from the lock screen, like a point-of-sale payment.

ABSTRACT

We identify and evaluate single-handed "dexterous gestures" to physically manipulate a phone using the fine motor skills of fingers. Four manipulations are defined: shift, spin (yaw axis), rotate (roll axis) and flip (pitch axis), with a formative survey showing all except flip have been performed for various reasons. A controlled experiment examines the speed, behaviour, and preference of manipulations in the form of dexterous gestures, by considering two directions and two movement magnitudes. Results show rotate is rated as easiest and most comfortable, while flip is rated lowest. Using a heuristic recognizer for spin, rotate, and flip, a one-week usability experiment finds increased practice and familiarity improve the speed and comfort of dexterous gestures. Design guidelines are developed to consider comfort, ability, and confidence when mapping dexterous gestures to interactions, and demonstrations show how such gestures can be used in smartphone applications.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**.

KEYWORDS

Interaction techniques, Finger dexterity, Mobile input, Gesture recognition

ACM Reference Format:
Yen-Ting Yeh, Fabrice Matulic, and Daniel Vogel. 2023. Phone Sleight of Hand: Finger-Based Dexterous Gestures for Physical Interaction with Mobile Phones. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3544548.3581121

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581121

1 INTRODUCTION

Our hands are remarkable when one considers the diverse ways we can grasp and manipulate objects. Aristotle and Anaxagoras even argue that human intelligence evolved due to the capability of human hands [21, 31]. When we interact with mobile phones, we already use a range of hand functions: from thumb input while gripping the phone, to physical interactions like squeezing [33], shaking, and wrist rotation [13]. Researchers have also proposed more elaborate types of input using wrist rotation (e.g. [1, 34, 57]) and motion gestures performed with the arm and wrist (e.g. [38, 39]). However, interactions proposed so far primarily use a power grip [25] in which the hand firmly grasps the phone during interactions, typically requiring muscular strength and producing larger movements in space.

We investigate an under-explored type of physical interaction that uses the opposite of a static power grip: a dynamic precision grip enabled by finger dexterity. By definition, dexterous manipulations include gross movements like juggling, but we focus on those that are decoupled from arm movement, i.e. *in-hand dexterous manipulations* [20]. This category uses a loose grip to allow object position and orientation to be manipulated primarily using smaller finger muscles for finer movements. In general, people have phenomenal ability to develop finger dexterity skills for activities such as playing musical instruments, specialized tasks in industry and healthcare, and crafts like knitting [32], but it is unclear if this innate human ability could also be used for phone interaction.

There have been limited demonstrations of in-hand dexterous gestures for input devices. For example, Soap [2] is a custom pointing device using mid-air manipulative gestures to interact with large wall displays, and MagPen [16] enables the detection of different dexterous pen-spinning gestures. With phones, Eardley et al. [6] note how people using a phone with one hand will loosen their grip to shift it down with finger movements when reaching far targets. Recently, Yang et al. [54] elicited ways to switch between front and back screens on a "dual-display" phone, finding some people loosened their grip to turn the phone over using a series of finger movements with the same hand. These works further motivate a systematic investigation into in-hand dexterous gestures for phone interaction to answer the research question: *"What is a general class of dexterous gestures for phone manipulation that are usable and acceptable to users"*. We imagine using this new style of interaction to trigger global or contextual actions, such as silencing a call, activating a voice assistant, or triggering a point-of-sale payment app (Figure 1). Such in-hand dexterous gestures would be complementary to power grip motion gestures since finger-based rotational movements likely have different motion characteristics, and importantly, since they are not limited by wrist range-of-motion, they enable full phone rotations.

We examine four dexterous phone manipulations: *shifting* by moving the phone up or down by "walking" the fingers; *spinning* by pinching the phone with one finger and the thumb and spinning it with the other fingers or using gravity; *rotating* by rolling the phone inside a loose grip; and *flipping* by turning the phone end-over-end by swapping fingers on the front or back of the phone. We use a multi-step methodology to understand how well users can perform those manipulations, what users' perceptions and preferences are, whether such interactions can be recognized algorithmically, and what kind of applications can make use of them. A formative study establishes people have some familiarity with these manipulations, and a controlled experiment measures their performance as well as gathers data to train a heuristic recognizer for spinning, rotating, and flipping. Finally, a three-phase usability study examines these three gestures after one week of practice with real-time recognition, and looks at differences in usage context like sitting versus standing. Our results show people can perform all types of gestures, with good recognition for spinning, rotating, and flipping. Rotating is fastest and most preferred, then spinning and flipping, with speed and acceptance increasing for all gestures with practice.

To summarize, we make three contributions: (1) a formal identification of in-hand dexterous manipulation as a novel class of physical phone interaction; (2) empirical evidence that a subset of in-hand dexterous gestures are practical in terms of user preference and performance; and (3) demonstrations showing in-hand dexterous gestures can be recognized reliably and used for a variety of practical applications.

2 RELATED WORK

Our work relates to the general area of single-handed phone interaction and motion gestures. We focus on phones, but discuss other devices and contexts when considering previous applications of dexterity in HCI.

2.1 Motion Gestures

Motion gestures are interactions in which users intentionally move the device in space to issue commands. Different from touch input, motion gestures can be performed without visual feedback, easing demands on user attention.

Perhaps the simplest example is Hudson et al.'s whack gesture [15] where the palm or heel of the hand firmly strikes the phone. Using an elicitation study, Ruiz et al. [39] found users propose diverse motion gestures like shaking, rotating, quickly moving a phone back and forth, or moving a phone to a specific body location for command invocation. DoubleFlip [38] is a wrist gesture that rotates the phone away and back to trigger actions. Motion gestures can also use other forms of input for context, for example Hinckley and Song [13] explored motion gestures combined with touch, like shaking the phone while touching an icon to execute a contextual command. Yang et al. [54] used an elicitation study to examine methods to switch between front and back screens on dual-display phones. Among user proposals, there were single-handed motion gestures to turn the phone along the roll axis and a method to roll the phone in the hand using the fingers. The latter is an example of a dexterous manipulation which we study more generally.

Many phone input methods use forms of tilting as a kind of motion gesture, such as tilt-based gestures for navigating documents, menus, or lists [8, 12, 29, 35] and sharing files with other devices [9]. A well-known example was originally proposed by Hinckley et al., where tilting a phone to the side changes the interface orientation between portrait and landscape [12]. Even a limited set of smaller tilt movements can be expanded into a useful input vocabulary. For example, Baglioni et al. proposed eight quick back-and-forth gestures discriminated by device acceleration and tilt direction [1]. These previous works do not specifically discuss tilting actions performed only with fingers; in our observation, all demonstrate tilting with a fixed grip using wrist movement. Rahman et al. [34] analyzed how well people can control tilt angle along three axes of wrist movement. They found 12 levels can be accurately controlled along the flexion to extension direction and 16 levels along the pronation to supination direction with a quadratic control-display function for tilt angle. We use tilting as a relative comparison in our work, but we ask participants to perform it only using their fingers without significant wrist movement.

Our gestures also work on a commodity phone with IMU sensors, but instead of whacking or waving with gross motor skills, or making limited rotations with wrist-based movement, we explore a distinctly different interaction space when the phone is manipulated independently using finger dexterity. This enables a novel class of gestures not limited by the biomechanical constraints of wrist and arm movements as in previous work.

2.2 In-hand Manipulation

In-hand manipulations are a class of dexterous gestures when holding, moving, and manipulating an object with one hand. These are essential interactions used in daily activities such as writing with a pen or using chopsticks. ToolStone [36] is an input device that is rotated, flipped, and tilted using the non-dominant hand. Based on how the device contacts a tablet, different commands are triggered, like tool selection, 3D model navigation, and viewport selection. Van Laerhoven et al. [48] created a cube-shape device which can sense its orientation and movement with built-in accelerometers. Gestures such as shaking, twisting, and knocking were used for device control and navigation. Soap [2] is a pointing device created by placing an optical sensor core inside an elastic fabric hull. Although

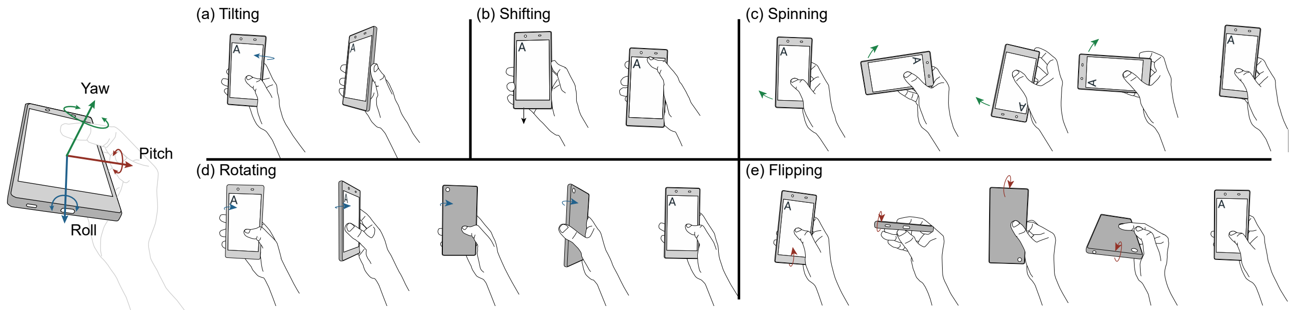

Figure 2: Types of dexterous finger manipulations with example variations: (a) *tilting*, showing a ‘tilt left’ variation, (b) *shifting*, showing a ‘shift down’ variation; (c) *spinning*, showing a ‘spin clockwise 360°’ variation; (d) *rotating*, showing a ‘rotate right 360°’ variation; (e) *flipping*, showing a ‘flip away 360°’ variation. Several manipulations can be described relative to the canonical pitch, yaw, and roll axes (shown at left).

the device was used to remotely control a large display, it included a dexterous gesture consisting in rotating the core 360° in the hand, which is similar to our rotating manipulation. To support active reading, Yoon et al. [58] detect tablet grips and motions, including a “lateral swing” gesture when the tablet is passed to another user via a combination of top grips and rotation of the tablet. This gesture has similar characteristics to our spinning manipulation.

Dexterous gestures have been explored extensively for pens, such as tilting to select menu items [45] or reveal layers [11], rolling to scroll web pages [44], switch modes [3], acquire buttons [43], undo activities [11], or rotate graphical objects [3, 11]. MagPen [16] can sense different pen-spinning and balancing gestures to trigger actions such as choosing ink properties and undoing strokes in a sketching application. Inspired by these dexterous techniques with other devices, we investigate extended forms of dexterous manipulations for use with a mobile phone.

In summary, previous work studied motion gestures with hand or wrist movements using phones, and dexterous object manipulations with pens or custom devices. Our work complements and significantly bridges these two spaces by exploring the new space of dexterous finger gestures for mobile phones.

3 DEXTEROUS MANIPULATIONS

Dexterous manipulations include a wide variety of actions. There are dexterous manipulations associated with moving objects in space, like those used in sports, magic acts, and circus performances. For example, juggling, tossing, twirling, and manipulating cards. An extreme application of this for phones was demonstrated in the ThrowMe phone app [26] in which a phone is tossed into the air to capture kinetic photos or bird’s-eye view images. Balanced spinning on a single support point is another form of dexterous object manipulation. Book, plate, or ball spinning can be seen in tricks and circus performances. A more common example is how people spin a pen on the side of their hand. Although unusual, balanced single-point phone spinning can be performed with excellent skills, as demonstrated in social media videos [59].

Compared to those somewhat acrobatic acts requiring fine motor skills, in-hand dexterous gestures combining finger movements with support of the palm are likely easier to perform and therefore more suitable for everyday use. Popular examples include using a fidget spinner [52] and manipulating Chinese “Baoding balls” for exercise and stress relief [51]. Ma and Dollar [20] studied this type of dexterous manipulations for the purpose of encoding human hand dexterity into robotic hands. They defined six primary in-hand manipulative movements: regrasping, in-grasp manipulation, finger gaiting, finger pivoting, rolling, and sliding. Regrasping is a movement that momentarily releases the object followed by a quick “regrasp” in a modified position or orientation. In-grasp manipulation is a movement to make small changes to the object’s orientation without removing the fingers. Finger gaiting is when the object is moved by replacing grasping fingers with free fingers in a cyclic alternating fashion. Finger pivoting is a manipulation while holding the object with two fingers and using other free fingers to rotate the object about the axis formed by the two finger points. Rolling is a movement to move the object by rotating it with a fixed pivot point. Sliding is a manipulation to move the object with a controlled slip. Our focus is on using these kinds of dexterous in-hand manipulations as explicit input for a phone. We define four new dexterous manipulations along with tilting, a simpler manipulation to use as a baseline (Figure 2):

Tilting is a type of “in-grasp manipulation” that changes the phone orientation similar to tilt-based interactions in previous work, but using only finger movements instead of the wrist. A typical sequence of finger motions is: grip the phone between the thumb, ring, and middle fingers, then use index and pinky fingers on the back or the top of the phone to tilt forward or backward; anchoring the side of the phone with middle, ring, and pinky fingers while moving the thumb up to tilt left; or anchoring the side of the phone with palm and thumb, then letting other fingers slide along the back of the phone to tilt right. Variations are defined using direction and angle, such as tilt right 90°, or tilt backward 45°.

Shifting translates the position of the phone relative to the palm along the *roll* axis. The motion typically uses the palm to support the phone while finger positions change in order to shift the device up or down in more than one step. A smaller shift can be achieved with “regrasping”, where the phone is pushed up or pulled down with the fingers in one movement. The up or down direction is used to define variations.

Spinning circles the phone around the *yaw* axis using a “finger pivoting” dexterous movement. It is performed by pinching the phone with the thumb on top and index or middle finger at the back, then typically using the free fingers to spin the phone. For smaller spins in certain directions, gravity alone can be relied upon for the movement once the phone is pinched. Spinning variations can be defined using direction (clockwise or counterclockwise), angle (e.g., 90° or quarter turn, 180° or half turn, and 360° or full turn), and speed (slow or fast).

Rotating is circling the phone around the *roll* axis in the palm using "rolling" and "sliding" dexterous movements. We define it as an extended movement of the more common left and right variations of tilting manipulations. At the end of a left or right tilt, the side of the phone slides along the bottom of the fingers and palm until the screen is against the palm. Then the grip is adjusted in a regrasping motion, with the action repeated if needed. Variations include direction (right or left), angle (e.g., 180° or half turn, and 360° or full turn) and speed (slow or fast).

Flipping is circling the phone around the *pitch* axis using a form of "finger gaiting" movement. We define it as an extended movement of the more common forward and backward variations of tilting manipulations. At the end of a forward or backward tilt, the thumb and fingers are swapped from the front and the back of the phone. This is repeated as necessary for larger movement angles. Variations include direction (forward or away), angle (e.g., 180° or half turn, and 360° or full turn) and speed (slow or fast).

3.1 Formative study

We conducted a formative study in the form of a questionnaire with self-guided tasks to understand previous experiences and preferences for the five types of dexterous manipulations defined above. We hypothesized that tilt is simpler and more familiar, so we treated it as a baseline to compare with the four more elaborate dexterous interactions. The questionnaire was divided into three parts: (1) demographic information including phone size and hand size, (2) previous experience with dexterous manipulations, and (3) preferences after trying each manipulation in a self-guided task[1].

3.1.1 Participants. We recruited 30 participants (19 males, 11 females) through flyers, word-of-mouth, and social media on a volunteer basis without remuneration. Most participants (28) reported their phone experience as more than 6 years of daily use. Participants used 17 different phone models with screens from 4.7 to 6.5 inches and 25 participants used phone cases. The circumference of the palm of the dominant hand (i.e. "glove size" [30]) ranged from 16.5 to 26.2 cm.

3.1.2 Procedure. Participants were asked to fill the questionnaire on a device other than their phone and to have their phone ready to try the manipulations.

In the experience part, each participant watched an animated demonstration of each manipulation and selected the ones they had done before, even if infrequently. For each manipulation they had previously experienced, participants were asked about frequency and reasons for doing them. For frequency, they were asked how often they performed the gesture on a daily, weekly, monthly, or less frequently than monthly basis. To explain why they performed a gesture, participants selected one or more reasons: reach specific location of the phone, change phone orientation, play games, fun, unintentional, and other.

In the tryout part, each manipulation was explained using a text description and animated demonstration similar to the previous part. Participants were instructed to hold their phone using a loose grip so that they could use the fingers of their dominant hand to manipulate the phone, with only the palm to support the device if necessary. Participants were asked to try to perform the manipulation shown in the animated demonstration, preferably over a soft surface such as a couch, a bed, or towels in order to avoid damaging their phone if accidentally dropped. The variations of manipulations tested were TILT in four directions with 45° to 90° magnitude, SHIFT up and down, SPIN, ROTATE, and FLIP in two directions with 360° magnitude. After each manipulation, participants were asked to rate their preference for ease and comfort on a 7-point Likert scale. The session required approximately 10 minutes.

3.1.3 Results. For previous experience, at least 4 participants had previous experience with all types of manipulations. SHIFT was the most common manipulation (86% of participants), followed by TILT (66%), SPIN (53%), ROTATE (43%), and FLIP (13%). For the easiness rating, most participants found all manipulations easy, except FLIP. TILT was considered the easiest movement with 91% of participants agreeing more or less strongly, followed by 88% for ROTATE, 71% for SHIFT, 61% for SPIN and 28% for FLIP. For the comfort rating, most participants found all manipulations except FLIP comfortable. TILT was considered the most comfortable gesture with 85% of participants agreeing more or less strongly, followed by 78% for ROTATE, 65% for SHIFT, 51% for SPIN, and 13% for FLIP.

3.2 Dexterous Gestures

With the manipulations defined above, dexterous gestures can be broken down into discrete atomic actions using specific combinations of manipulation variations, or continuous input of a parameter. For example, rotating 180° clockwise to decline an incoming call, or adjusting the volume based on the tilt angle. Due to phone size and physical hand motion constraints, tilt and shift manipulations are bound in their extent and repetition. However, spin, rotate, and flip manipulations can form unlimited gestures with infinite angles. Sequences of discrete actions can also form variations of dexterous gestures, including within manipulations (e.g., spin clockwise 90° then spin counterclockwise 90°), or between manipulations (e.g., flip 180° followed by rotate 180° and spin 180° to return original orientation).

We mainly focus on single discrete atomic actions to explore the gesture space in terms of people's previous experience with dexterous manipulations, user preference, gesture speed, reliability of gesture detection, and what applications are suitable for them.

4 EXPERIMENT 1: PERFORMANCE AND PREFERENCE

The results of the formative study demonstrated most dexterous manipulations were performed in the past by users for various reasons, and most were considered easy and comfortable. The goal of this experiment is to determine the speed of dexterous gestures, how participants perform them, and their preferences. Shift, spin, rotate, and flip manipulations were tested as dexterous gestures

[1]See supplementary materials for full study questionnaire and additional correlation analysis of phone weight, thickness, etc.

with two directions and two movement magnitudes. The study was conducted remotely due to constraints imposed by the Covid pandemic, so participants performed the designated gestures with their own phone. We measured the time to complete each gesture, collected internal sensor data, and recorded reasons for incomplete gestures as well as subjective preferences.

4.0.1 Participants. Participants were recruited using flyers and word-of-mouth, and received a $25 remuneration for their participation. Participants were required to have full use of their right hand and fingers and have access to an Android phone with built-in IMU sensors. From the total 26 participants who completed the experiment, 8 were removed after examining their data: 2 had missing sensor data, 2 had gesture-ending detection errors, and 4 appeared to have not followed the experiment procedure, as revealed by almost "flat" sensor data with no obvious movement, or almost identical sensor patterns for some gestures. The remaining 18 participants completed the experiment successfully, 9 females and 9 males, with average age 26.8 years (SD = 4.0). Smartphone experience, phone characteristics, and hand size were recorded as in the formative study (summarized in Table 1).

4.0.2 Apparatus. The experiment was deployed as an Android 6.0+ app APK. Data from accelerometer, magnetometer, gyroscope, light, and proximity sensors were logged with a 50 Hz update rate. Touch input location, size, and pressure were also logged. Each trial was recorded to a file then uploaded to cloud storage. Participants were asked to remove any accessories other than protective phone cases and set the phone to "do not disturb" to avoid interruptions. The app executed in portrait orientation with auto-rotate disabled.

4.0.3 Task. Before starting each trial, an illustration of the gesture (similar to Figure 2) and an animated demonstration of each gesture was shown (Figure 3a). Each trial began by tapping a start button with the right thumb. The size and position of the button were such that it was comfortable to reach with a normal grip. Next, participants were asked to hold the phone still with their normal gripping posture (Figure 3b) for one second until a beep sounds. A simple visualization of the phone's movement was shown to help participants get a feel for the threshold according to which the device was considered still. After the beep, they started performing the gesture using only their fingers. At the end of the gesture, they were told to hold the phone still for one second again and waited for another beep. After a second beep, they returned to the normal grip posture and began the next trial. Participants were allowed to use their other hand to help return the phone to the start position between trials.

If the participant believed they performed the gesture incompletely or incorrectly, they pressed a "redo" button, provided a reason for the failure, and then repeated the trial again. The possible reasons were "phone dropped", "discontinuous movement", "app interrupted", "unfinished movement", or "other".

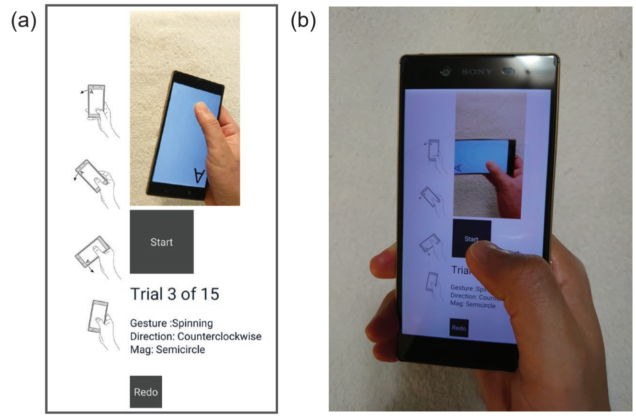

Figure 3: Experiment 1 task: (a) the interface before starting a trial; (b) the phone was held in the dominant hand.

4.0.4 Procedure. The experiment was divided into three parts: pre-session instruction; main session with measured trials; and post-experiment questionnaire. Each participant attended a 10-minute one-on-one online meeting with instructions and a question-and-answer period. During this time, the participant installed the Android app and verified it was working as expected, the flow of the experiment was introduced, the task explained, and general guidance for completing the study was given. Next, the participant went through the main session of measured trials covering each gesture at a convenient time for them. This main session lasted approximately 45 minutes. Participants were asked to be seated and to do the experiment on top of a soft surface (e.g. bed, couch) or use towels in case they accidentally dropped their phone. Additionally, they were requested not to rest their forearm or hand on any supporting surface. After the main session was completed, participants rated each gesture on four aspects using a numeric rating from 1 to 7: ease, comfort, confidence, and social acceptance. The experiment was approximately one hour in total. The full text of the questions is provided in the supplementary material[1].

4.0.5 Design. Our study follows a within-subjects design with three primary independent variables: MANIPULATION { SHIFT, SPIN, ROTATE, FLIP }; DIRECTION { ADD, ABD }; and for all MANIPULATION conditions except SHIFT, MAGNITUDE { HALF, FULL }. ADD is adduction and describes gestures toward the middle of the body: shift down, spin counterclockwise, rotate right, and flip forward. ABD is abduction and describes gestures away from the body: shift up, spin clockwise, rotate left, and flip away. This creates 14 different

Table 1: Experiment 1 demographics (18 participants in total).

Smartphone experience (years)		Daily phone usage (hours)		Phone size (inches)		Hand size (mm)	
3-5	1	Less than 1	1	4-5	1	139-165	2
6-10	13	1-2	4	5-6	11	165-190	5
More then 10	4	2-4	6	More than 6	6	190-215	4
		4-8	6			215-241	5
		More than 8	1			241-266	2

dexterous gesture conditions, one for each combination of MANIPULATION, DIRECTION, and MAGNITUDE. Each participant completed 15 trials, including two practice ones, for each gesture condition as one sequence, with the order of all gesture conditions randomized. In summary: we recorded 182 completed trials per participant, 3276 trials in total.

There are five dependent measures: *Time* is the time from the start (the first beep) until the end (one second before the second beep) of the gesture. *Ease*, *Comfort*, *Confidence*, and *Acceptance* are numeric ratings for ease-of-use, comfort, confidence of not dropping the phone, and willingness to do the gesture in public.

4.0.6 Analysis. We used the $1.5 \cdot$ IQR (interquartile range) rule to detect trial outliers for each combination of participants, MANIPULATION, DIRECTION, and MAGNITUDE according to trial time. In total, 216 trials (6.6%) were removed. Due to the unbalanced design for SHIFT without MAGNITUDE, to analyze the effect of MANIPULATION, a MANIPULATION × DIRECTION ANOVA with Bonferroni-corrected pairwise comparisons was used. To understand the effect of MAGNITUDE, we removed the trials of SHIFT and used a MANIPULATION × DIRECTION × MAGNITUDE ANOVA with Bonferroni-corrected pairwise comparisons. Residuals for *Time* were not normally distributed, so Tukey's Ladder of Powers transformation [46] was used for statistical analysis. We visually inspected the Q-Q plot to confirm normality. Aligned Rank Transform [53] was used for numeric ratings as the distribution was not normal. Figure 4 summarizes main results for dexterous gesture conditions with a summary by MANIPULATION. Spearman correlation tests were used for the phone form factor and hand size analysis. We focused on phone size for simplicity since it normally correlates with other factors such as weight, height, width, and thickness. The full table for all phone factor correlation results can be found in the supplementary materials[1], and analysis scripts can be obtained on a public repository[2].

4.1 Results

To streamline the presentation of results, details of statistical tests and significant differences are provided as tables in the Appendix. References are in the form "A.1: Table 1a" where A.1 refers to subsection 1 of the Appendix.

4.1.1 Time. We found ROTATE is the fastest gesture and about 0.4s, 0.7s, and 0.9s faster than SHIFT, FLIP and SPIN (Figure 4a; see A.1: Table 3a(i) for statistical tests showing MANIPULATION main effect). The mean time for ABD, and ADD are 2.98s and 3.08s respectively (but no significant main effect of DIRECTION). For SHIFT, movement in the ADD direction is faster than ABD; and for FLIP, movement in the ABD direction is faster (see A.1: Table 4a for statistical tests showing MANIPULATION and DIRECTION interaction). Overall, HALF gestures are 0.96s faster than FULL (see A.1: Table 3b(i) for statistical tests showing MAGNITUDE main effect). For FULL gestures, ROTATE is the fastest manipulation in both directions (see A.1: Table 4b for specific pairwise differences showing MANIPULATION, DIRECTION, and MAGNITUDE interaction). In addition, participants with larger hands can perform dexterous gestures slightly quicker (Spearman correlation showed a negative weak relationship between gesture time and hand size (r(3058) = -0.16, $p < .001$)).

[2]https://github.com/exii-uw/phone-dexterity

4.1.2 Ease. We found ROTATE was considered the easiest gesture and FLIP the least easy; HALF gestures were, as expected, rated easier than FULL gestures (Figure 4b; see A.1: Table 3a(ii) and b(ii) for statistical tests showing MANIPULATION and MAGNITUDE main effect, but no interaction effect). The ease rating is lower when performing the gestures with a larger phone (Spearman correlation showed a negative weak relationship between ease and phone size (r(250) = -0.19, $p < .01$) and also between ease and hand size (r(250) = -0.18, $p < .01$)).

4.1.3 Comfort. Participants considered ROTATE the most comfortable gesture and FLIP the least; and HALF gestures were considered more comfortable than FULL (Figure 4c; see A.1: Table 3a(iii) and b(iii) for statistical tests showing MANIPULATION and MAGNITUDE main effect, but no interaction effect). The comfort rating is lower when performing the gestures with a larger phone (Spearman correlation showed a negative weak relationship between comfort and phone size (r(250) = -0.15, $p < .05$) and also between comfort and hand size (r(250) = -0.17, $p < .01$)).

4.1.4 Confidence. We found participants are most confident about not dropping their phone for ROTATE and SHIFT, and least confident with FLIP, but all ratings were neutral or above (Figure 4d). Participants also have higher confidence in HALF gestures than FULL (see A.1: Table 3a(iv) and b(iv) for statistical tests showing MANIPULATION and MAGNITUDE main effect, but no interaction effect). Participants with smaller phones tend to be more confident performing the gestures (Spearman correlation showed a negative weak relationship between confidence and phone size (r(250) = -0.27, $p < .001$) and moderate relationship between confidence and hand size (r(250) = -0.44, $p < .001$)).

4.1.5 Social Acceptance. We found SHIFT and ROTATE are the gestures that participants are most willing to perform in front of people or in public areas (Figure 4e). They also perceive HALF gestures are more socially acceptable than FULL gestures (see A.1: Table 3a(v) and b(v) for statistical tests showing MANIPULATION and MAGNITUDE main effect, but no interaction effect). Using dexterous gestures in public was deemed more acceptable with a smaller phone (Spearman correlation showed a negative weak relationship between acceptance and phone size (r(250) = -0.16, $p < .01$) and also between acceptance and hand size (r(250) = -0.17, $p < .01$)).

4.2 Summary

Overall, rotating is the fastest manipulation with the highest rating for ease and comfort. Shifting is rated as more socially acceptable, which may be due to it also being the most familiar manipulation. However, compared to rotating, the ease and comfort score is lower, likely because of the loosened grip and the relative difficulty of the gesture. This result is similar to Eardley et al.'s findings [6], where loosening and shifting grips were associated with lower comfort and secure scores. Spinning is considered slower, especially with full magnitude. This is likely because the gesture includes a short shifting movement between each half spin. Flip gestures are the least preferred for ease, comfort, and confidence. However, a half flip away (abduction) gesture can be performed in 2.16s, which is comparable to the fastest gesture times. The movement of this gesture is similar to pen-spinning techniques, which may be the reason

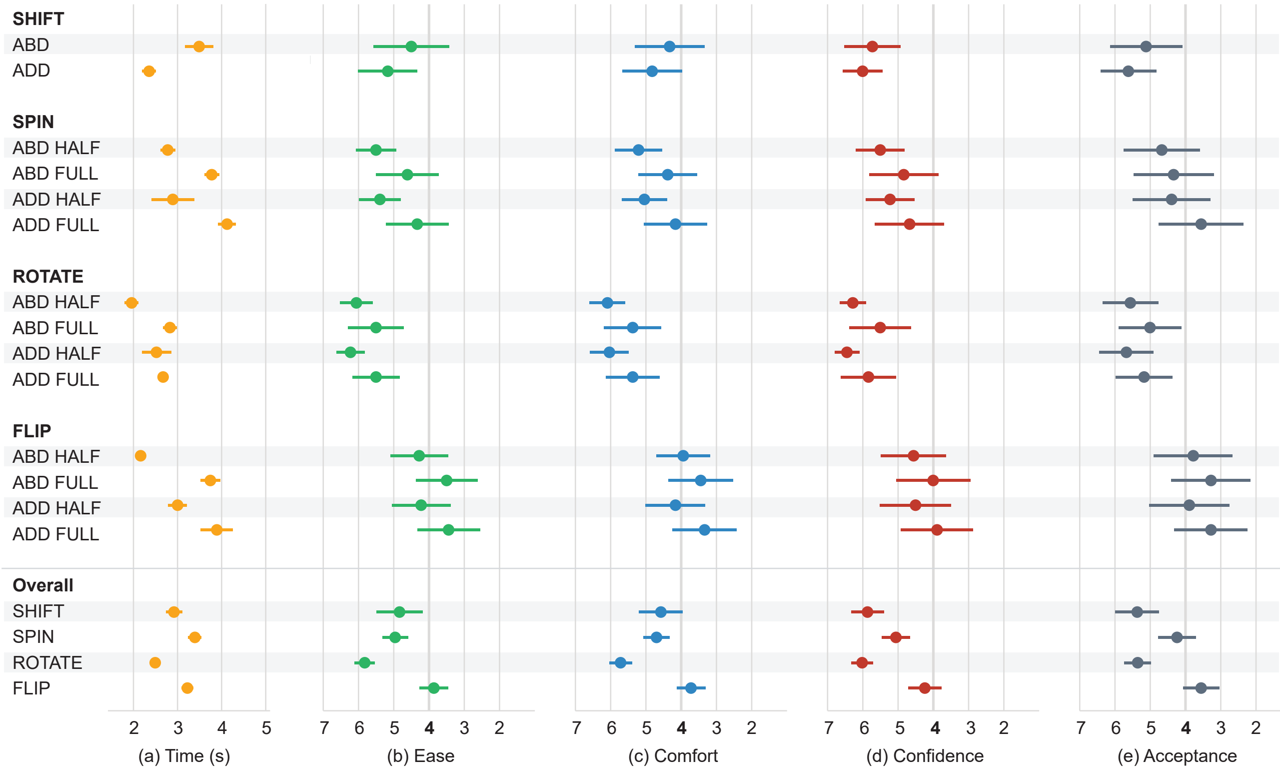

Figure 4: Comparison of gestures defined by MANIPULATION, DIRECTION, MAGNITUDE, as well as overall by MANIPULATION: (a) Time; (b) Ease rating; (c) Comfort rating; (d) Confidence rating; (e) Social Acceptance rating. Note: rating scales inverted to enable comparison with time, left-most points in each sub-graph are better. (error bars are 95% confidence intervals)

for its speed. In summary, rotate is perhaps the most promising manipulation, especially rotate gestures with half magnitude.

Due to our experiment protocol, our results for the gesture time may not exactly be representative of real use. In order to record clean and complete sensor data while the gesture was performed, participants were asked to hold the phone still at the start and end and wait for a beep sound. This filtered out extraneous movements such as lifting the thumb after pressing the start button, but likely added some reaction time. Additionally, gesture times recorded in controlled conditions might be different from in-the-wild gestures detected using motion thresholds. Furthermore, since dexterity skills are typically learned and honed through time, experiments over longer periods of time would be needed to examine possible learning effects and determine how fast users can ultimately execute such dexterous manipulations.

To obtain a better understanding of confidence and ability, we examined the "redo reasons" when participants did not complete trials. From all the possible reasons, phone dropped was the most critical issue since the consequence was possible phone damage. Within the whole experiment (3,276 completed trials), 24 redos were due to phone drops, including 13 times during flipping, 6 times during shifting, 4 times during spinning, and once during rotating. Notably, 15 cases of such phone drops occurred in the first 7 trials. This is somewhat consistent with the lowest confidence score given to flip manipulations. Combining phone dropped and discontinuous movement as reasons for redoing a trial, 31 such "redos" happened during flipping, followed by rotating (27), spinning (25), and shifting (21). It is possible that unwanted contacts between screen and fingers or the palm may cause standard system gestures to be triggered, causing the experiment app to be interrupted. Another possibility is pressing the power button during a gesture. However, we found these were not too frequent. Only 12 such interruptions during flip, 8 during rotate, 9 during spin, and 4 during shift were recorded.

5 PROTOTYPE SYSTEM

We create an Android prototype with a rule-based recognizer based on sensor patterns in order to demonstrate the potential of dexterous gestures. Examining the IMU sensor data, we can identify patterns for different gestures such as the z value of the accelerometer dropping from near 1 to almost -1 when rotating or flipping 180°, and the x, y, or z value from the gyroscope sensor mainly affected by the axis corresponding to flipping, rotating, or spinning gestures. Based on a visual comparison with IMU patterns in the H-MOG dataset of phone use while reading or texting [42], the sensor pattern of shift gestures are likely very hard to distinguish from normal movements. We made several attempts to recognize shift gestures using deep learning methods, including LSTM and CNN models, but those interactions could not be reliably discriminated.

Consequently, we focus on flipping, rotating, and spinning for the recognizer and applications below.

5.1 Recognizer

We build a recognizer based on accumulated differences of quaternions (which are commonly used to represent rotations) to classify spinning, rotating, and flipping with two directions and magnitudes. Differences of quaternions simplify the raw IMU sensor data so that rotation angles can be better distinguished. We sum the quaternion differences between each consecutive frame of each rotation sensor axis to compute the angle difference, and check the value of the corresponding Euler axis for different gestures (x for flip, y for rotate, and z for spin). For example, for a quaternion representing a rotation θ around the z-axis during a spin gesture, the quaternion difference on the z-axis would be $\sin\frac{\theta}{2}$. With a high sampling frame rate, the accumulated quaternion difference on the z-axis for a half spin gesture would approach $\frac{\pi}{2}$.

$$\lim_{k\to\infty} k \sin\frac{\pi}{2k} = \frac{\pi}{2} \tag{1}$$

To reduce false positives, we only recognize a gesture when three conditions are satisfied for accelerometer data: (1) at least one axis has a zero crossing, (2) at least two axes cross each other, and (3) at least one axis has a difference greater than 3 m/s^2.

A sliding time window segments real-time data for recognition. The average time for FULL gestures in the experiment data is 3.52s (SD = 1.74) and 2.56s (SD = 2.13) for HALF gestures. We therefore use a 4s window for FULL and 2.5s window for HALF gestures. The sensor update rate is 50 Hz, and the system checks for a gesture every 0.2s. For each check, we calculate the accumulated quaternion difference and see if a FULL gesture was performed within the 4s window, or a HALF gesture within the 2.5s window. Since a FULL gesture includes a HALF gesture, we introduce an additional 0.2s delay after recognizing a HALF gesture to test if it actually ended, or if the phone is still rotating to perform a FULL gesture. This means the maximum delay for recognizing a gesture action is 0.4s.

5.1.1 Threshold Selection. We analyzed the data collected from the experiment to determine thresholds to detect each dexterous gesture. There are 2615 trials after removing shift gestures and 193 outliers (6.9%) using the same $1.5 \cdot \text{IQR}$ method as the experiment. To further improve consistency, we also applied the same $1.5 \cdot \text{IQR}$ to identify outliers for each MAGNITUDE according to accumulated quaternion differences, which removed another 348 trials (12.4%). With this dataset, we found the average accumulated quaternion difference in the corresponding Euler axis for HALF gestures across participants is 1.4 (SD=0.2), and 2.68 (SD=0.44) for FULL gestures. They are approximately equivalent to 160° (SD = 22°) and 308° (SD = 50°) Euler angles. This indicates that participants tend to rotate less than expected, so a lower angle detection threshold is needed to conform to actual user behaviour.

To fine-tune those thresholds, we tested our recognizer on a 6-person (10%) subset of the Extrasensory [47] dataset of in-the-wild phone usage (210 hours). Like DoubleFlip [38], we used the rate of false positives per 8 hours as our metric. Figure 5a shows ROC curves plotting average accuracy using our dataset and maximum false positive rates across all 12 gestures. The five curves plot different threshold combinations. To minimize false positives and maximize accuracy, we choose thresholds with accuracy higher than 75%, and false positive less than 3. The selected thresholds are 1.3 for HALF and 2.02 for FULL gestures, which are approximately equivalent to 149° and 231° Euler angles.

5.1.2 False Positive Test with Datasets. We tested our recognizer on two datasets: H-MOG [42] (341 hrs of more stable phone usage) and Extrasensory [47] data not used for threshold selection (54 people, 1514 hrs of in-the-wild usage with more diverse movements).

Figure 5b shows the rate of false positives of each gesture per 8 hours. With H-MOG, adduction spinning has a higher rate (1.24 full, 0.91 half), likely due to similarity with landscape and portrait changes. All other gesture rates are less than 0.28. With Extrasensory, half-rotations have higher rates (1.89 abduction, 1.91 adduction). We believe this is likely due to movements when setting down or picking up the phone. The rates for the other two half gestures are low: SPIN (0.91 abduction, 0.75 adduction) and FLIP (1.01 abduction, 0.45 adduction); and all full gestures are below 0.59. For comparison, the single DoubleFlip gesture has a rate of one false positive per 8 hours [38].

5.1.3 True Positive Test with Users. To evaluate recognition accuracy in real-time, we recruited 12 participants: 6 females and 6 males, average age of 25.8 years (SD = 2.8). Five also participated in our previous experiment conducted more than 11 months before. The apparatus, task, and procedure are similar to Experiment 1, but with the addition of the gesture recognizer and 12 dexterous gesture conditions (i.e. without shifting). Participants completed 2 practice trials and 5 measurement trials for each gesture condition as one sequence, 60 measurement trials per participant The dependent measure is the recognizer accuracy.

Overall, our recognizer shows high accuracy: ROTATE (97.9%), FLIP (91.7%), and SPIN (85%) (Figure 5c). For specific dexterous gestures, both ROTATE-ABD, ROTATE-ADD-HALF, and FLIP-ADD-FULL were recognized perfectly. SPIN-HALF had the lowest accuracy (71.7% for ABD and 76.7% for ADD), likely due to participants sometimes stopping a spin gesture early when the accumulated quaternion difference had not reached the required threshold. This can happen after the phone contacts the palm.

5.1.4 Limitations and Improvements. Our recognizer based on quaternion differences cannot distinguish gestures that are performed only with fingers from similar phone movements using the wrist. However, due to anatomical constraints, it is not possible to perform full gestures or half spins using only the wrist. For half gestures, additional sensor data could distinguish those actions. For example, wrist manipulations with power grip tend to not touch the screen while finger-based dexterous gestures do. The threshold selection plays a critical role in our recognizer, especially for reducing the false positives. Selecting thresholds for individual manipulations could address those with higher false positives, such as choosing a higher half threshold for rotate gestures.

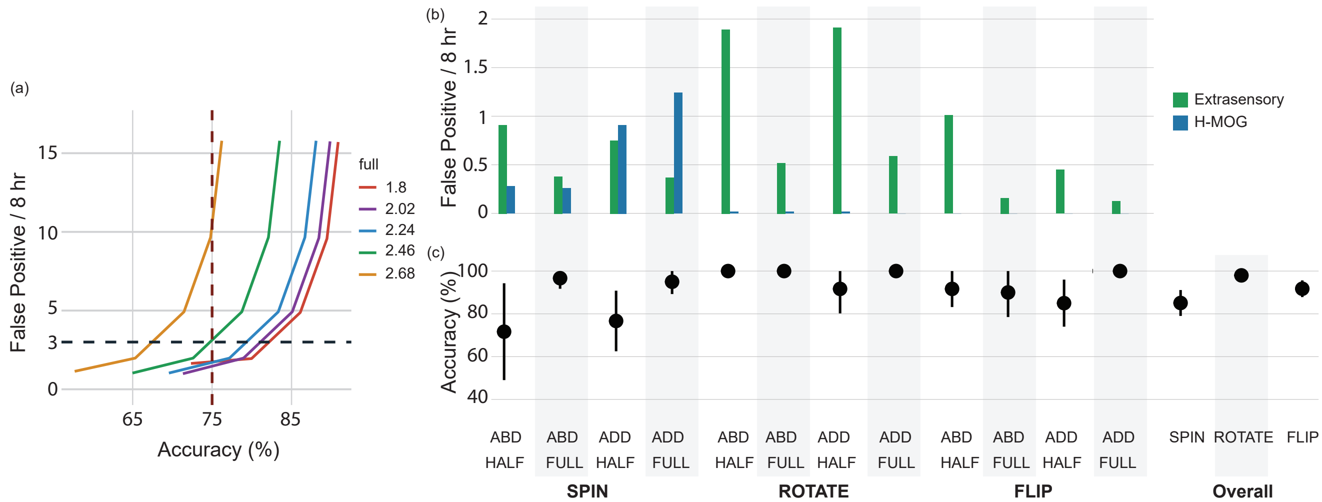

Figure 5: Recognizer validations: (a) ROC curves of average accuracy and maximum false positive rate among all gestures: for 25 different FULL and HALF thresholds, each curve shows different HALF thresholds (1.4, 1.3, 1.2, 1.1, and 1.0) when used with the same FULL threshold; (b) false positive analysis of Extrasensory and H-MOG dataset; (c) accuracy of gestures as defined by combination of MANIPULATION, DIRECTION, MAGNITUDE, as well as overall by MANIPULATION (error bars are 95% confidence).

5.2 Applications

We consider potential applications making use of dexterous gestures, for which there are general design principles and constraints[3]:

- Gestures ending with the screen away from the user are only useful for tasks that require no immediate visual feedback with only little touch input. Such gestures would be suitable for voice input and output.
- Gestures inverting the phone so the microphone is up and close to the mouth are useful for voice commands [55].
- Less preferred and more cumbersome gestures are more suitable for infrequent commands, or commands that incur a high penalty if triggered accidentally (e.g. power off, system diagnose).
- Gestures should preferably be activated from the lock screen to minimize accidental touches.

Dexterous gestures can be used as global commands (e.g., opening a camera app while the phone is locked, invoking assistance tools, or checking time and weather) or interaction with applications (e.g., declining an incoming call, dismissing an alarm, or issuing commands to a music player). We implement applications in the following categories (please also see the accompanying video for demonstrations):

Functions Without Visual Input or Feedback — Dexterous gestures can be performed without looking at the screen or visual feedback. This can be useful when the phone screen is not immediately visible, such as when it is in a bag. An example scenario is declining a phone call by reaching into the bag and rotating the phone (Figure 1a). The risks of dropping and damaging the phone are significantly reduced when the device is in a bag, which may lower the barrier for using dexterous manipulations in such situations. Gestures can also benefit users with visual impairments. For blind people, dexterous gestures expand the input options they have to quickly and conveniently trigger phone actions [49].

[3]Content in "5.2 Applications" also appears in our UIST 2022 demonstration [56].

Application Shortcuts — Opening particular apps and looping between opened apps with pre-defined gestures enable simple and direct commands with or without visual feedback. For example, rotating left full to open a calendar, and spinning clockwise half to open a mobile payment application (Figure 1c). Spinning can be used to loop through or swap opened apps. Spinning clockwise or counterclockwise full could switch to the next or previous app. Although the flip gesture might be more difficult to perform, it can be used to open infrequent but critical apps, such as flipping away full to open system settings and flipping forward full to power off.

Camera — The rear or front camera can be opened by rotating left or right full directly without unlocking the device. Rotating left half would open the rear camera with auto capturing, or users could tap the screen to take a photo as back-of-device interaction.

Voice Notes and Intelligent Assistant Queries — ProxiTalk [55] showed that bringing the phone to the mouth is a promising method to activate speech input. A half flip of the phone can bring the microphone up to record audio. The flip away gesture can be used for dictating voice notes (Figure 1b), and flip toward gesture could open the search function. The phone can be rotated right half to hear time and weather information.

Alarm Functions — Using fine motor skills to perform dexterous gestures requires concentration, which can reduce the likelihood of unintentional operations [19]. For example, rotating right full can dismiss an alarm, or rotating left half and full can snooze the alarm 5 and 15 minutes respectively, instead of using swipe gestures.

Music player — Dexterous gestures can also be mapped to functions inside an application like a music player. A full rotate could change the song and a half rotate could skip forward or backward. A half spin can control the volume while a full spin can mute or un-mute the phone directly. Because rotating gestures can be performed in a narrow space, changing songs with rotating gestures in the pocket may be useful while running or training.

6 EXPERIMENT 2: PRACTICE AND CONTEXT

We conducted a one-week experiment to further examine the performance, preference, and usability of dexterous gestures after users gain more familiarity and practice. Because a half gesture is included in a full gesture, we focus on three "full" manipulations for spin, rotate, and flip, each in two directions. To examine usage context, sitting and standing conditions were tested. Participants used their own phone throughout the study.

6.0.1 Participants. Participants were recruited using our institution's student mailing list and word-of-mouth, each received \$50 for completing the study. With the same phone requirements as Experiment 1, we recruited 12 participants, ages 23 to 31 (M=26.83, SD=2.79), of which 8 were male, and 4 were female. Note that 4 of these participants also completed Experiment 1 more than one year before. Smartphone experience, phone characteristics, and hand size are summarized in Table 2.

6.0.2 Dexterity Training App. We created a dexterity training app that detects each gesture, counts the repetitions, and displays scores for smoothness and speed in a graphical style reminiscent of meditation apps[4]. Users can track their progress in terms of gesture speed and smoothness over multiple days. The scores are calculated according to the deviation of quaternion differences between frames and gesture time, and the app displays simple graphical rewards when thresholds of these scores are exceeded. The idea is that the app encourages users to manipulate the phone smoothly and quickly, and also trains the dexterity of fingers (similar to Chinese "Baoding balls" [51]). Source code is available on the project's public repository[2].

6.0.3 Procedure. The experiment was conducted in three phases: pre-practice, practice, and post-practice.

The *pre-practice* phase was conducted in-person. An experiment app was installed on the participant's phone similar to the one used in the true positives experiment (Section 5.1.3). After receiving instructions about the 6 dexterous gestures and experiment task, participants completed measured trials while sitting. At the end of the session, they provided subjective ratings and then installed a second app for dexterity training.

For the *practice* phase, the participant used the training app at home for at least 10 minutes every day for 7 days.

The *post-practice* phase was conducted after practice. Five participants completed it in-person 1 day after completing practice, and the rest completed it remotely using a live video call 7 to 9 days after practice. There were two post-practice sections: first, participants completed the same measured trials as those in pre-practice while sitting and also when standing. Then, they answered additional questions about their preferences in multiple scenarios, and their feedback about demonstrations and possible applications was recorded[1].

[4]The training app is demonstrated in the accompanying video.

6.0.4 Design. We used a within subjects design with three primary independent variables: SESSION with 2 levels (BEFORE, AFTER practice); MANIPULATION with 3 levels (SPIN, ROTATE, FLIP); and DIRECTION with 2 levels (ADD, ABD). There was another independent variable for the AFTER practice condition: SCENARIO with 2 levels (SIT, STAND). We tested STAND in the AFTER practice condition to understand the performance and preference of gestures in a more difficult scenario. As such, there are a total of 18 gesture conditions: (12 SESSION × MANIPULATION × DIRECTION + 6 MANIPULATION × DIRECTION for AFTER & STAND). There were 7 trials per gesture condition, including two practice ones. The order for SESSION was fixed, the order for SCENARIO was counter-balanced using a Latin square, and the order for MANIPULATION × DIRECTION was randomized. In summary we recorded 90 completed trials per participants, i.e. 1080 trials in total.

The primary measures obtained or computed from logs are *Accuracy*, *Time*, and *Smoothness*. *Accuracy* is the gesture accuracy of our proposed recognizer. *Time* is the gesture time from the start of the trial until the gesture is recognized. *Smoothness* is calculated from the quaternion difference in continuous frames while the gesture is executed. We define high smoothness using two criteria: (1) the quaternion difference values of the corresponding Euler axis for different gestures should be roughly constant, and (2) the quaternion difference values of the other two axes should be close to 0. Specifically, a gesture generates a series of accumulated quaternion difference values $[QD_1...QD_n]$. Each QD_i has components representing the three Euler axes: X_i, Y_i, and Z_i. We calculate *Smoothness* as the sum of two terms: (1) the sum of absolute differences between each primary axis component with the median primary axis component, and (2) the sum of the components for the other two axes. For example, Y is the primary axis for the rotate gesture, so *Smoothness* is calculated as:

$$Smoothness_{rotate} = \frac{\sum_{i=1}^{n} |Y_i - Mdn(Y)|}{n} + \frac{\sum_{i=1}^{n} |X_i| + |Z_i|}{2 \times n} \quad (2)$$

There are four subjective measurements for each dexterous gestures which are the same as in the previous experiment.

6.0.5 Analysis. To analyze the effect of SESSION, we remove the trials of STAND and use a SESSION × MANIPULATION × DIRECTION ANOVA with Bonferroni-corrected pairwise comparisons. To understand the effect of SCENARIO, we remove the trials of BEFORE practice and use a SCENARIO × MANIPULATION × DIRECTION ANOVA with Bonferroni-corrected pairwise comparisons. Greenhouse-Geisser correction is used when there is a sphericity violation. We use generalized linear mixed models for *Accuracy* analysis because the distribution is close to a Poisson distribution. Residuals for *Time*

Table 2: Experiment 2 demographics (12 participants in total).

Smartphone experience (years)		Daily phone usage (hours)		Phone size (inches)		Hand size (mm)	
6-10	5	1-2	3	5-6	3	165-190	2
More then 10	7	2-4	4	More than 6	9	190-215	2
		4-8	5			215-241	5
						241-266	3

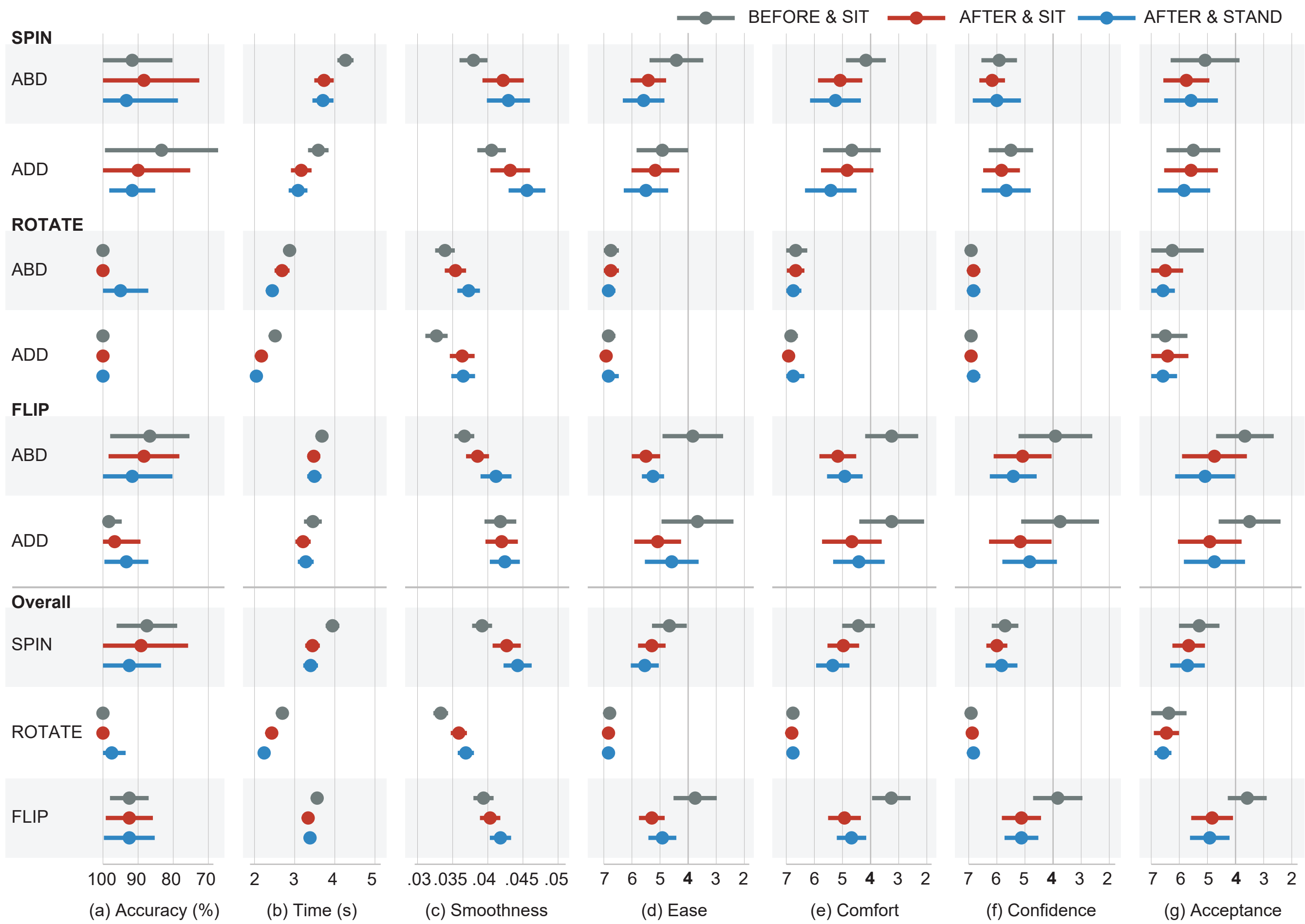

Figure 6: Comparison of gestures defined by MANIPULATION, DIRECTION, as well as overall by MANIPULATION: (a) Accuracy; (b) Time; (c) Smoothness; (d) Ease rating; (e) Comfort rating; (f) Confidence rating; (g) Social Acceptance rating. Note: rating scales inverted to enable comparison with time, left-most points in each sub-graph are better. (error bars are 95% confidence intervals)

and *Smoothness* are not normally distributed, so Tukey's Ladder of Powers transformation [46] is used. Aligned Rank Transform [7, 53] is used for numeric ratings due to a non-normal distribution. Figure 6 summarizes the main results for dexterous gesture conditions with a breakdown by MANIPULATION type.

6.1 Results for Before and After Practice

We only report the main effect of SESSION, or interactions involving SESSION. For *Accuracy* and *Smoothness*, there was no effect.

Time — Participants can perform dexterous gestures 0.3s faster AFTER practice (Figure 6b; see A.2: Table 5a(i) for statistical tests showing SESSION main effect).

Ease, Comfort, Confidence, and Social Acceptance — Participants rated all four subjective scores higher AFTER practice than BEFORE (Figure 6d, e, f, g; see A.2: Table 5a(ii, iii, iv, v) for statistical tests showing SESSION main effect). FLIP is rated easier, more comfortable, and more socially acceptable AFTER practice (see A.2: Table 5b(i, ii, iii, iv) for statistical tests showing SESSION and MANIPULATION interaction).

6.2 Results for Sitting versus Standing

There were no main effects or interactions involving SCENARIO, so we only report main effects for MANIPULATION and DIRECTION.

Accuracy — There was no effect of SCENARIO, MANIPULATION, and DIRECTION. Overall, our recognizer has high accuracy: 94% for both SIT and STAND (Figure 6a).

Time — ROTATE is 1.2s and 1.0s faster than SPIN and FLIP (Figure 6b; see A.3: Table 6a(i) and b(i) for statistical tests showing MANIPULATION and DIRECTION main effect).

Smoothness — ROTATE is better than SPIN and FLIP (Figure 6c; see A.3: Table 6a(ii) for statistical tests showing MANIPULATION main effect).

Ease, Comfort, Confidence, and Social Acceptance — We found ROTATE was rated highest in all four subjective ratings, and FLIP received the lowest ratings for confidence and acceptance (Figure 6d, e, f, g; see A.3: Table 6a(iii, iv, v, vi) for statistical tests showing MANIPULATION main effect).

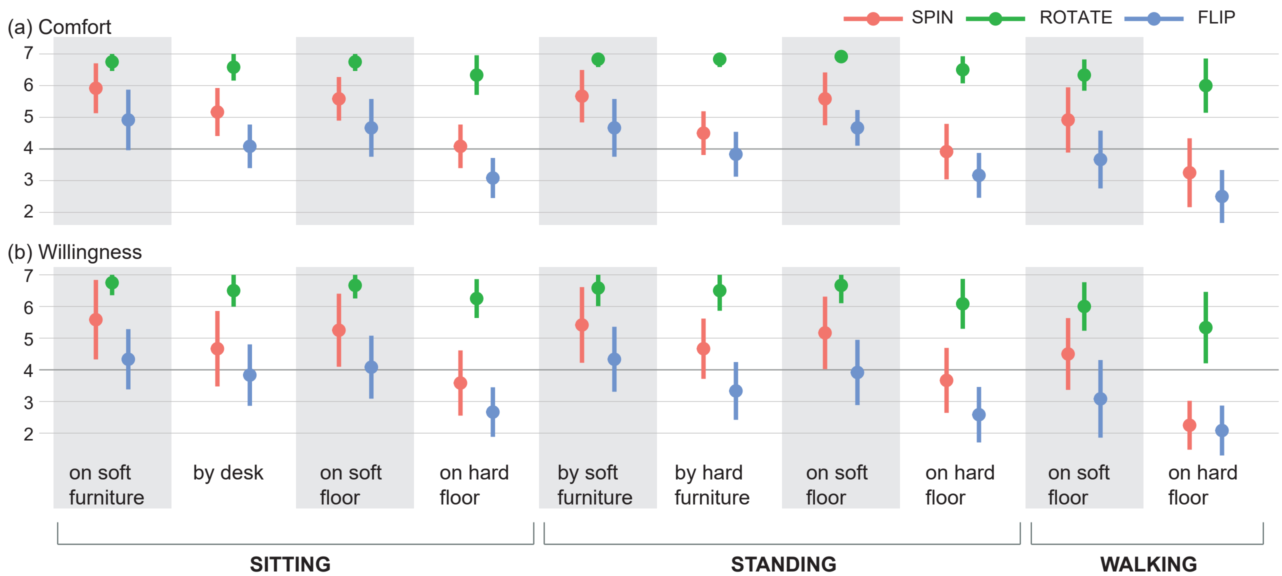

Figure 7: Comparison of scenarios: (a) *Comfort* rating; (b) *Willingness* rating. (error bars are 95% confidence intervals)

6.3 Questionnaire Results

To better understand comfort and willingness to perform dexterous gestures in multiple scenarios, two additional subjective measures are included in the questionnaire at the end of the post-practice session. The scenarios are split into sitting, standing, and walking in various environments. Figure 7 summarizes the main results for performing dexterous gestures in different contexts. Most participants found ROTATE comfortable and were willing to perform it in all situations, even while walking on hard floor (comfort: 6, willingness: 5.3). Participants expressed they were more comfortable performing dexterous gestures while sitting on, standing by, or walking on soft surfaces such as soft furniture (e.g. couch, bed) or floor (e.g. grass). Following the lower comfort ratings, participants were less willing to flip the phone especially on a hard surface or while walking. However, some participants were more willing to do half-flips, e.g. *"I like the half flip, not the full one. I can do it quite comfortably."* [P10]

To gain more insights about how participants would be willing to perform dexterous gestures practically in their daily lives, they watched the video of demonstrations (section 5.2) and provided feedback as well as proposed other applications. Most participants found the demonstrations practical and expressed they would like to use them, especially for snoozing and dismissing alarm, or opening the payment application: *"I would definitely use the half spinning to open the payment app. It's easy and I don't need to open the NFC manually."* [P4]; *"Snoozing an alarm for a certain period based on a gesture is really nice."* [P5]; and *"I like the half flip to take notes because the flipping gesture is more controllable than the voice input."* [P7]. On the other hand, participants also reported that they would not use dexterous gestures compared to current phone gestures, such as pressing buttons, or shaking phone: *"I can shake my phone to open the flashlight in Motorola Moto G5+, so I prefer that than rotating."* [P7]; and *"If I were already using the phone (i.e. phone is awake), swiping or tapping are easier than dexterous gestures."* [P2]

Participants made some interesting suggestions for potential applications making use of dexterous gestures, such as making an emergency call, integrating gestures in games to increase interactivity, or helping hand rehabilitation: *"The gestures would be useful in the situations that making a movement without letting other people know, like calling the police with a simple rotate, or starting phone recording directly."* [P9]; *"The gestures can be applied into games to increase the fun elements, such as flipping the phone to fire weapons."* [P3]; and *"...using such gestures with the training app to rehabilitate people who are suffering partial disabilities in their hands due to a stroke or injury because the smoothness and speed scores are good indicators of improvement!"* [P5]

6.4 Summary

Overall, the results for the pre-practice session align with experiment 1 and the true positive test with users in section 5.1.3: all full gestures can be recognized accurately at a rate above 88%; rotating is the fastest manipulation with the highest subjective ratings; spinning is slower; and flipping is rated lowest. After one week of practice, the speed and subjective ratings of dexterous gestures improved, especially the comfort and confidence ratings. Some participants found better ways to perform gestures during practice: *"For spinning gestures, I found the sweet spot to pinch the phone, and used gravity and momentum to spin the phone quickly"* [P10]. Most participants became more confident about not dropping their phone: *"I become more comfortable and confident to do the gestures, even while talking to other people"* [P3].

We found some evidence of a trade-off between gesture speed and smoothness. Spearman correlation showed a negative moderate relationship in spinning ($r(358) = -0.58$, $p < .001$), rotating ($r(358) = -0.47$, $p < .001$), and flipping ($r(358) = -0.52$, $p < .001$). To increase smoothness, participants seem to slow down for more control, e.g. *"I found that doing the gesture slowly can increase the score for spinning"* [P11].

There was no quantitative differences between sitting and standing. Some participants did comment about feeling less confident, *"... while standing I noticed that I was more careful trying not to drop my phone"* [P5], but others felt more comfortable because of the increased range of motion, *"I prefer standing because the arm can fall down naturally, I have to lift up my arm to hold the phone while sitting"* [P7].

7 DISCUSSION

We discuss and summarize design recommendations based on overall results.

7.0.1 Manipulation. The formative study and the experiments show that the rotating manipulation is perceived as the easiest and most comfortable with higher social acceptance, which suggests it can be used to trigger regular phone actions. The rotation half gesture recognition has higher false positives, likely because it is accidentally triggered when putting down or picking up the phone. This makes it less suitable for global commands, but we believe it can still be used for contextual functions such as declining an incoming call, dismissing an alarm, or triggering functions in an active application. Spinning the phone may require more time and finger movement, but this is a familiar gesture that was rated as easy. Spinning gestures are well suited for functions which need visual feedback since the screen remains visible during the movement. As exemplified by Yoon et al.'s lateral swing gesture [58], spinning manipulations can be used in collaborative situations like sharing content with coworkers. Since this motion involves several people, it could be used to protect privacy, such as locking the phone with a spin. Flipping gestures should be used less frequently since they had lower ratings and were associated with higher chances of dropping the phone. However, half flip gestures, especially away from the body (abduction), are relatively fast and rated high enough to warrant use for less frequent functions.

7.0.2 Magnitude. Rotating and flipping halfway end with the phone screen facing away. This means these gestuers should be used to trigger functions that do not require visual feedback or touch input, such as using speech and audio for note dictation, and dismissing a call. Recent commercial developments suggest phones with screens on both sides could become more common [22, 23, 40]. The practical benefits of half gestures are more evident for dual screen phones as a way to switch between screens [54]. These explicit motions would be distinguishable from simple static detection of phone orientation to trigger specific actions. For example, users could switch between main and secondary screens to view multiple applications using half-rotation gestures, or display private content using half flips.

Full gestures can be improved after practice. With increased speed and comfort, performing full gestures to activate the camera, turn on the flashlight, or start a recording can be useful with current phones. One advantage of gestures relying on finger dexterity rather than full wrist or arm motions is that they can be repeated indefinitely. Due to hand anatomy limitations, only half gestures can be performed with the fingers in a power grip. With a loose grip and dexterous finger manipulations, multiple phone rotations are possible. Although the speed of such gestures would be slower, they can be used to control a continuous parameter such as increasing the duration of the alarm snooze by rotating the phone multiple times. Individual dexterous gestures can also be combined for security purposes, like unlocking the device after 2 full-right rotations, 1 full-left rotation and 3 full-flip-away gestures.

7.0.3 Accidental Input. Accidental input when performing dexterous gestures, such as touching the screen with the palm or pressing the power button while moving fingers, may be a concern. This only happened a few times in our experiment (1%), but it still is something to be addressed for reliability. Methods such as recognizing palm touch events [17], detecting unintentional touch events similar to palm rejection for pen input [41] or grip recognition [18] can be applied to reject accidental inputs. Restricting dexterous gestures to the lock screen would also largely mitigate this issue.

7.0.4 Single-hand vs Two-hand Gesturing. Our interaction space is defined by in-hand manipulations, so we only examined single-handed gestures. Single-hand phone usage is important for phone interaction techniques since the other hand may be encumbered [28] and people use their phone more often with one hand than two [14]. However, single-hand dexterous gestures can also be performed with some assistance of the other hand. For example, flipping the phone with fingers on both sides to make sure the weight of the device is equally distributed and grip stability is increased. Users may wish to first safely practise their dexterous gesturing skills using two hands before perfecting them with one hand. These aspects, as well as learning effects, can be explored in future work.

7.0.5 Risks. Although our results showed that people could perform dexterous gestures when holding the phone in a loose grip, there were a few cases of phone drops, especially with the flipping gesture. But with some practice, those risks diminish as users gain more confidence. When running the studies, we asked participants to perform gestures above a soft surface. This may have lead to higher subjective scores compared to other "riskier" situations where the gestures are performed while standing or walking on a hard floor, which is shown in the results of the questionnaire. However, phone protection accessories such as rubber cases and screen protectors, may help reduce user apprehension by alleviating the risk of phone damage from accidental drops. A more thorough examination of these situations is required to obtain a better understanding of benefits versus risks.

7.0.6 Fatigue. Large motion gestures may cause "gorilla arm" [4, 10], but this kind of fatigue is unlikely with dexterous gestures since the arm can remain at a comfortable position. However, dexterous movements require a high amount of finger movement, which likely introduces muscle fatigue in the hand. In our experiments, participants could take breaks between blocks or pause practising when they felt finger or hand soreness. We found they usually required a break after multiple blocks, but generally felt comfortable performing single manipulations, especially after a full week of practice. This suggests that applications requiring many dexterous gestures during a concentrated time should be avoided. For this reason, most of our applications demonstrate dexterous gestures for less frequent, single manipulations. Future work could specifically investigate fatigue in dexterous gestures, perhaps over a longer period or in a controlled way where the number of gestures per time span is controlled.

7.0.7 Practical Usage Verification. Although we collected ideas for how dexterous gestures could be used in experiment 2, our participant feedback was based on our demonstration videos and their imagination. Future work should explore and validate how practical these potential applications are.

7.0.8 Comparisons with Conventional Gestures. We did not conduct experiments to compare dexterous gestures with standard phone interaction techniques for two reasons. First, dexterous gestures are complementary to other forms of phone input like touch, squeezing, and motion gestures: our ultimate goal is to increase expressiveness with phones, not to replace current methods. Second, the goal of this work is to gain an understanding of dexterous gestures, how usable and socially acceptable they are, whether they can be reliably recognized, what kind of applications could exploit them. We recognize that dexterous gestures appear novel to most users, and by definition, they require an element of skill to perform. For example, it is likely that simple gestures, such as swiping, tapping, and even squeezing, would be rated as faster and easier to perform. Below, we offer some high-level comparisons with other phone input techniques with respect to speed and diversity of gesture set, memorability and semantic mapping, and eyes-free interaction.

Dexterous gestures can be used as direct commands with comparable speed to methods combining a delimiter and subsequent commands [27]. All 12 dexterous gestures can be reliably detected with very high true positive rates and low false positive rates. These rates could likely be further improved by optimizing our recognizer. Addtionally, the top-speed of half-gestures is about 2 seconds and 3 seconds for full-gestures after practice. Consider how DoubleFlip [38] and Active Edge [33] are single gestures used to delimit a subsequent action to specify the actual command. With a greater diversity in our dexterous gesture set, we can directly trigger multiple different commands. In terms of speed, dexterous gestures are comparable to using the DoubleFlip motion gesture to delimit a command mode with a flick motion (average 3.22 s) [27].

The action of some dexterous gestures can have matching semantic associations to improve their memorability [24]. For instance, the spinning gesture performs a lateral rotation which suggests giving or sharing, and therefore could be associated with payment or file sending actions. Flip brings the microphone up and close to the mouth, which creates a possible association with voice commands. Using longer full-gestures makes sense for prolonged actions, such as snoozing an alarm for a longer time.

Dexterous gestures also lend themselves to eyes-free interaction. Negulescu et al. [27] found that motion gestures can decrease the time looking at the smartphone during walking, and since dexterous gestures require even less motion, that finding likely applies as well. A very promising application of eyes-free dexterous gesturing is for people with visual impairments [50]. In an elicitation study, Dim and Ren [5] found that motion gestures are more efficient for blind users, but Romano et al. [37] found that blind users used motion gestures less often because they were unfamiliar and concerned about accidentally hitting nearby objects. Dexterous gestures may have an advantage because they are highly tactile when learning and they require no large movement of the hand or arm.

8 CONCLUSION

We explored a new form of physical phone interactions called dexterous gestures which use fine motor skills of fingers to manipulate the device in-hand. We defined a gesture design space consisting of shifting, spinning, rotating, and flipping manipulations, with tilting used as a baseline. A formative study showed that all manipulations except flipping had been previously performed by participants. A performance experiment showed that rotating was fast and the most preferred gesture while a full flip was rated lowest. A prototype system using a heuristic recognizer demonstrated that most spinning, rotating, and flipping gestures can be recognized reliably on standard phones with 91.2% average accuracy, which illustrates how this style of gestures could be used in real applications. A one-week experiment further showed that speed and willingness to adopt dexterous gestures improve after practising, and that there is little difference in using the gestures while sitting or standing. Our exploration shows how human dexterity can be harnessed for new forms of phone interaction.

ACKNOWLEDGMENTS

This work made possible by the NSERC Discovery Grant 2018-05187, and the Canada Foundation for Innovation Infrastructure Fund 33151 "Facility for Fully Interactive Physio-digital Spaces".

REFERENCES

[1] Mathias Baglioni, Eric Lecolinet, and Yves Guiard. 2011. JerkTilts: Using Accelerometers for Eight-Choice Selection on Mobile Devices. In *Proceedings of the 13th International Conference on Multimodal Interfaces* (Alicante, Spain) *(ICMI '11)*. Association for Computing Machinery, New York, NY, USA, 121–128. https://doi.org/10.1145/2070481.2070503

[2] Patrick Baudisch, Mike Sinclair, and Andrew Wilson. 2006. Soap: A Pointing Device That Works in Mid-Air. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology* (Montreux, Switzerland) *(UIST '06)*. Association for Computing Machinery, New York, NY, USA, 43–46. https://doi.org/10.1145/1166253.1166261

[3] 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* (Monterey, CA, USA) *(UIST '08)*. Association for Computing Machinery, New York, NY, USA, 191–200. https://doi.org/10.1145/1449715.1449745

[4] 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* (Melbourne, Australia) *(OZCHI '09)*. Association for Computing Machinery, New York, NY, USA, 161–168. https://doi.org/10.1145/1738826.1738853

[5] Nem Khan Dim and Xiangshi Ren. 2014. Designing Motion Gesture Interfaces in Mobile Phones for Blind People. *Journal of Computer Science and Technology* 29, 5 (01 Sep 2014), 812–824. https://doi.org/10.1007/s11390-014-1470-5

[6] Rachel Eardley, Anne Roudaut, Steve Gill, and Stephen J. Thompson. 2017. Understanding Grip Shifts: How Form Factors Impact Hand Movements on Mobile Phones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4680–4691. https://doi.org/10.1145/3025453.3025835

[7] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 754–768. https://doi.org/10.1145/3472749.3474784

[8] Beverly L. Harrison, Kenneth P. Fishkin, Anuj Gujar, Carlos Mochon, and Roy Want. 1998. Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Los Angeles, California, USA) *(CHI '98)*. ACM Press/Addison-Wesley Publishing Co., USA, 17–24. https://doi.org/10.1145/274644.274647

[9] Nabeel Hassan, Md. Mahfuzur Rahman, Pourang Irani, and Peter Graham. 2009. Chucking: A One-Handed Document Sharing Technique. In *Human-Computer Interaction – INTERACT 2009*. Springer Berlin Heidelberg, Berlin, Heidelberg, 264–278. https://doi.org/10.1007/978-3-642-03658-3_33

[10] 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* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[11] Ken Hinckley, Xiang 'Anthony' Chen, and Hrvoje Benko. 2013. Motion and Context Sensing Techniques for Pen Computing. In *Proceedings of Graphics Interface 2013* (Regina, Sascatchewan, Canada) *(GI '13)*. Canadian Information Processing Society, CAN, 71–78. https://dl.acm.org/doi/10.5555/2532129.2532143

[12] Ken Hinckley, Jeff Pierce, Mike Sinclair, and Eric Horvitz. 2000. Sensing Techniques for Mobile Interaction. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology* (San Diego, California, USA) *(UIST '00)*. Association for Computing Machinery, New York, NY, USA, 91–100. https://doi.org/10.1145/354401.354417

[13] 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* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 801–810. https://doi.org/10.1145/1978942.1979059

[14] Steven Hoober. 2013. How Do Users Really Hold Mobile Devices? *UXmatters: Insights and inspiration for the user experience community* (18 Feb 2013). https://www.uxmatters.com/mt/archives/2013/02/how-do-users-really-hold-mobile-devices.php

[15] Scott E. Hudson, Chris Harrison, Beverly L. Harrison, and Anthony LaMarca. 2010. Whack Gestures: Inexact and Inattentive Interaction with Mobile Devices. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction* (Cambridge, Massachusetts, USA) *(TEI '10)*. Association for Computing Machinery, New York, NY, USA, 109–112. https://doi.org/10.1145/1709886.1709906

[16] Sungjae Hwang, Andrea Bianchi, Myungwook Ahn, and Kwangyun Wohn. 2013. MagPen: Magnetically Driven Pen Interactions on and around Conventional Smartphones. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Munich, Germany) *(MobileHCI '13)*. Association for Computing Machinery, New York, NY, USA, 412–415. https://doi.org/10.1145/2493190.2493194

[17] Huy Viet Le, Thomas Kosch, Patrick Bader, Sven Mayer, and Niels Henze. 2018. PalmTouch: Using the Palm as an Additional Input Modality on Commodity Smartphones. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173934

[18] Huy Viet Le, Sven Mayer, and Niels Henze. 2018. InfiniTouch: Finger-Aware Interaction on Fully Touch Sensitive Smartphones. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 779–792. https://doi.org/10.1145/3242587.3242605

[19] Christof Lutteroth, Moiz Penkar, and Gerald Weber. 2015. Gaze vs. Mouse: A Fast and Accurate Gaze-Only Click Alternative. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 385–394. https://doi.org/10.1145/2807442.2807461

[20] Raymond R. Ma and Aaron M. Dollar. 2011. On dexterity and dexterous manipulation. In *2011 15th International Conference on Advanced Robotics (ICAR)* (Tallinn, Estonia). IEEE, New York, NY, USA, 1–7. https://doi.org/10.1109/ICAR.2011.6088576

[21] Pablo Maurette. 2018. The Children of Anaxagoras. https://www.laphamsquarterly.org/roundtable/children-anaxagoras.

[22] mi.com. 2019. mi-mix-alpha. https://www.mi.com/global/mi-mix-alpha.

[23] microsoft.com. 2021. Surface Dou 2. https://www.microsoft.com/en-ca/d/surface-duo-2/9408kgxp4xjl.

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

[25] J. R. Napier. 1956. THE PREHENSILE MOVEMENTS OF THE HUMAN HAND. *The Journal of Bone and Joint Surgery. British volume* 38-B, 4 (1956), 902–913. https://doi.org/10.1302/0301-620X.38B4.902 PMID: 13376678.

[26] nbcnews.com. 2012. ThrowMeApp: The most fun you'll have destroying your Android. https://www.nbcnews.com/tech/gadgets/throwmeapp-most-fun-youll-have-destroying-your-android-flna1c6656834.

[27] Matei Negulescu, Jaime Ruiz, Yang Li, and Edward Lank. 2012. Tap, Swipe, or Move: Attentional Demands for Distracted Smartphone Input. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Capri Island, Italy) *(AVI '12)*. Association for Computing Machinery, New York, NY, USA, 173–180. https://doi.org/10.1145/2254556.2254589

[28] Alexander Ng, Stephen A. Brewster, and John H. Williamson. 2014. Investigating the Effects of Encumbrance on One- and Two- Handed Interactions with Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1981–1990. https://doi.org/10.1145/2556288.2557312

[29] Ian Oakley and Sile O'Modhrain. 2005. Tilt to scroll: evaluating a motion based vibrotactile mobile interface. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference* (Pisa, Italy). IEEE, New York, NY, USA, 40–49. https://doi.org/10.1109/WHC.2005.138

[30] palmflex.com. 2020. Glove Sizing. https://www.palmflex.com/glove-sizing.html.

[31] Carlos E. Perez. 2018. How Hand Dexterity leads to Human General Intelligence. https://medium.com/intuitionmachine/hand-dexterity-and-human-general-intelligence-af5596e1d36d.

[32] Stephen Pheasant and Christine M Haslegrave. 2005. *Bodyspace: Anthropometry, Ergonomics and the Design of Work*. CRC Press.

[33] Philip Quinn, Seungyon Claire Lee, Melissa Barnhart, and Shumin Zhai. 2019. Active Edge: Designing Squeeze Gestures for the Google Pixel 2. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 274:1–274:13. https://doi.org/10.1145/3290605.3300504

[34] Mahfuz Rahman, Sean Gustafson, Pourang Irani, and Sriram Subramanian. 2009. Tilt Techniques: Investigating the Dexterity of Wrist-Based Input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1943–1952. https://doi.org/10.1145/1518701.1518997

[35] Jun Rekimoto. 1996. Tilting Operations for Small Screen Interfaces. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology* (Seattle, Washington, USA) *(UIST '96)*. Association for Computing Machinery, New York, NY, USA, 167–168. https://doi.org/10.1145/237091.237115

[36] Jun Rekimoto and Eduardo Sciammarella. 2000. ToolStone: Effective Use of the Physical Manipulation Vocabularies of Input Devices. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology* (San Diego, California, USA) *(UIST '00)*. Association for Computing Machinery, New York, NY, USA, 109–117. https://doi.org/10.1145/354401.354421

[37] Marco Romano, Andrea Bellucci, and Ignacio Aedo. 2015. Understanding Touch and Motion Gestures for Blind People on Mobile Devices. In *Human-Computer Interaction – INTERACT 2015*. Springer International Publishing, Cham, 38–46. https://doi.org/10.1007/978-3-319-22701-6_3

[38] 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* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2717–2720. https://doi.org/10.1145/1978942.1979341

[39] 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* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/1978942.1978971

[40] samsung.com. 2019. galaxy-fold. https://www.samsung.com/ca/smartphones/galaxy-fold/.

[41] Julia Schwarz, Robert Xiao, Jennifer Mankoff, Scott E. Hudson, and Chris Harrison. 2014. Probabilistic Palm Rejection Using Spatiotemporal Touch Features and Iterative Classification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2009–2012. https://doi.org/10.1145/2556288.2557056

[42] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani. 2016. HMOG: New Behavioral Biometric Features for Continuous Authentication of Smartphone Users. *IEEE Transactions on Information Forensics and Security* 11, 5 (2016), 877–892. https://doi.org/10.1109/TIFS.2015.2506542

[43] Hyunyoung Song, Hrvoje Benko, Francois Guimbretiere, Shahram Izadi, Xiang Cao, and Ken Hinckley. 2011. Grips and Gestures on a Multi-Touch Pen. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1323–1332. https://doi.org/10.1145/1978942.1979138

[44] Yu Suzuki, Kazuo Misue, and Jiro Tanaka. 2007. Stylus Enhancement to Enrich Interaction with Computers. In *Proceedings of the 12th International Conference on Human-Computer Interaction: Interaction Platforms and Techniques* (Beijing, China) *(HCI'07)*. Springer-Verlag, Berlin, Heidelberg, 133–142. https://doi.org/10.1007/978-3-540-73107-8_15

[45] Feng Tian, Xiang Ao, Hongan Wang, Vidya Setlur, and Guozhong Dai. 2007. The Tilt Cursor: Enhancing Stimulus-Response Compatibility by Providing 3d Orientation Cue of Pen. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 303–306. https://doi.org/10.1145/1240624.1240675

[46] John W Tukey. 1977. *Exploratory data analysis*. Vol. 2. Reading, MA.

[47] Yonatan Vaizman, Katherine Ellis, and Gert Lanckriet. 2017. Recognizing Detailed Human Context in the Wild from Smartphones and Smartwatches. *IEEE Pervasive*

Computing 16, 4 (2017), 62–74. https://doi.org/10.1109/MPRV.2017.3971131
[48] Kristof Van Laerhoven, Nicolas Villar, Albrecht Schmidt, Gerd Kortuem, and Hans Gellersen. 2003. Using an Autonomous Cube for Basic Navigation and Input. In *Proceedings of the 5th International Conference on Multimodal Interfaces* (Vancouver, British Columbia, Canada) *(ICMI '03)*. Association for Computing Machinery, New York, NY, USA, 203–210. https://doi.org/10.1145/958432.958472
[49] Radu-Daniel Vatavu. 2017. Visual Impairments and Mobile Touchscreen Interaction: State-of-the-Art, Causes of Visual Impairment, and Design Guidelines. *International Journal of Human–Computer Interaction* 33, 6 (2017), 486–509. https://doi.org/10.1080/10447318.2017.1279827
[50] Radu-Daniel Vatavu and Jean Vanderdonckt. 2020. What Gestures Do Users with Visual Impairments Prefer to Interact with Smart Devices? And How Much We Know About It. In *Companion Publication of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS' 20 Companion)*. Association for Computing Machinery, New York, NY, USA, 85–90. https://doi.org/10.1145/3393914.3395896
[51] wikipedia.org. 2020. Baoding balls. https://en.wikipedia.org/wiki/Baoding_balls.
[52] wikipedia.org. 2020. Fidget spinner. https://en.wikipedia.org/wiki/Fidget_spinner.
[53] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963
[54] Zhican Yang, Chun Yu, Xin Chen, Jingjia Luo, and Yuanchun Shi. 2022. Investigating user-defined flipping gestures for dual-display phones. *International Journal of Human-Computer Studies* 163 (2022), 102800. https://doi.org/10.1016/j.ijhcs.2022.102800
[55] Zhican Yang, Chun Yu, Fengshi Zheng, and Yuanchun Shi. 2019. ProxiTalk: Activate Speech Input by Bringing Smartphone to the Mouth. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 118 (sep 2019), 25 pages. https://doi.org/10.1145/3351276
[56] Yen-Ting Yeh, Fabrice Matulic, and Daniel Vogel. 2022. Demonstrating Finger-Based Dexterous Phone Gestures. In *Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22 Adjunct)*. Association for Computing Machinery, New York, NY, USA, Article 69, 3 pages. https://doi.org/10.1145/3526114.3558645
[57] Hui-Shyong Yeo, Xiao-Shen Phang, Steven J. Castellucci, Per Ola Kristensson, and Aaron Quigley. 2017. Investigating Tilt-Based Gesture Keyboard Entry for Single-Handed Text Entry on Large Devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4194–4202. https://doi.org/10.1145/3025453.3025520
[58] Dongwook Yoon, Ken Hinckley, Hrvoje Benko, François Guimbretière, Pourang Irani, Michel Pahud, and Marcel Gavriliu. 2015. Sensing Tablet Grasp + Micro-Mobility for Active Reading. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 477–487. https://doi.org/10.1145/2807442.2807510
[59] youtube.com. 2018. How to Spin Your Phone on Your Finger! (Step by Step Tutorial). https://www.youtube.com/watch?v=_4evgE5Q0Go.

APPENDIX: TABLES OF STATISTICAL TESTS

This appendix presents tables of ANOVA and post hoc statistical tests for main effects and interactions of our results in experiment 1 and 2 (Section 4.1, 6.1, and 6.2).

A.1 Experiment 1: Results

Table 3: Main effect

(a) MANIPULATION

		(i) *Time*		(ii) *Ease*		(iii) *Comfort*		(iv) *Confidence*		(v) *Acceptance*	
		$F_{3,51} = 7.31$, $p < .001$, $\eta^2_G = 0.08$		$F_{3,51} = 29.52$, $p < .001$		$F_{3,51} = 30.93$, $p < .001$		$F_{3,51} = 29.78$, $p < .001$		$F_{3,51} = 21.22$, $p < .001$	
comparisons		*diff (s)*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*
ROTATE	SHIFT	-0.43	$< .001^{***}$	0.99	$< .01^{**}$	1.16	$< .001^{***}$	0.15	1	-0.01	1
ROTATE	SPIN	-0.9	$< .001^{***}$	0.86	$< .001^{***}$	1.03	$< .001^{***}$	0.95	$< .001^{***}$	1.11	$< .001^{***}$
ROTATE	FLIP	-0.73	$< .001^{***}$	1.96	$< .001^{***}$	2.02	$< .001^{***}$	1.77	$< .001^{***}$	1.79	$< .001^{***}$
SHIFT	SPIN	-0.47	$< .001^{***}$	-0.13	1	-0.13	1	0.8	$< .01^{**}$	1.12	$< .01^{**}$
SHIFT	FLIP	-0.3	$< .001^{***}$	0.97	$< .001^{***}$	0.86	$< .01^{**}$	1.62	$< .001^{***}$	1.8	$< .001^{***}$
SPIN	FLIP	0.17	.15	1.1	$< .001^{***}$	0.99	$< .001^{***}$	0.82	$< .01^{**}$	0.68	.05

(b) MAGNITUDE

		(i) *Time*		(ii) *Ease*		(iii) *Comfort*		(iv) *Confidence*		(v) *Acceptance*	
		$F_{1,17} = 36.42$, $p < .001$, $\eta^2_G = 0.19$		$F_{1,17} = 0.34$, $p < .001$		$F_{1,17} = 28.10$, $p < .001$		$F_{1,17} = 13.86$, $p < .001$		$F_{1,17} = 12.57$, $p < .001$	
comparisons		*diff (s)*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*
HALF	FULL	-0.96	$< .001^{***}$	0.8	$< .001^{***}$	0.74	$< .001^{***}$	0.63	$< .001^{***}$	0.56	$< .001^{***}$

Table 4: Interaction for *time*. Note: Only the comparisons with significant difference are shown.

(a) MANIPULATION × DIRECTION ($F_{3,51} = 10.69, p < .001, \eta^2_G = 0.05$)			
comparisons for ABD		*diff (s)*	*p-value*
ROTATE	SHIFT	-1.11	< .001***
ROTATE	SPIN	-0.89	< .001***
ROTATE	FLIP	-0.6	< .001***
FLIP	SPIN	-0.29	< .05*
comparisons for ADD		*diff (s)*	*p-value*
SHIFT	SPIN	-1.16	< .001***
SHIFT	FLIP	-1.11	< .001***
ROTATE	SPIN	-0.91	< .001***
ROTATE	FLIP	-0.86	< .001***
comparisons for SHIFT		*diff (s)*	*p-value*
ADD	ABD	-1.14	< .001***
comparisons for FLIP		*diff (s)*	*p-value*
ABD	ADD	-0.48	< .05*
(b) MANIPULATION × DIRECTION × MAGNITUDE ($F_{2,34} = 5.42, p < .01, \eta^2_G = 0.013$)			
comparisons for HALF		*diff (s)*	*p-value*
ROTATE-ABD	SPIN-ABD	-0.82	< .001***
ROTATE-ABD	SPIN-ADD	-0.93	< .001***
ROTATE-ABD	ROTATE-ADD	-0.57	< .05*
ROTATE-ABD	FLIP-ADD	-1.05	< .001***
ROTATE-ADD	SPIN-ABD	-0.25	< .01**
ROTATE-ADD	FLIP-ADD	-0.48	< .001***
FLIP-ABD	SPIN-ABD	-0.62	< .001***
FLIP-ABD	FLIP-ADD	-0.85	< .001***
comparisons for FULL		*diff (s)*	*p-value*
ROTATE-ABD	SPIN-ABD	-0.94	< .001***
ROTATE-ABD	SPIN-ADD	-1.3	< .001***
ROTATE-ABD	FLIP-ABD	-0.92	< .001***
ROTATE-ABD	FLIP-ADD	-1.06	< .001***
ROTATE-ADD	SPIN-ABD	-1.1	< .001***
ROTATE-ADD	SPIN-ADD	-1.46	< .001***
ROTATE-ADD	FLIP-ABD	-1.08	< .001***
ROTATE-ADD	FLIP-ADD	-1.22	< .001***

A.2 Experiment 2: Results for Before and After Practice

Table 5: Main effect and interaction for SESSION. Note: Only measures with significant difference are shown.

(a) SESSION

		(i) *Time* $F_{1,11} = 9.31$, $p < .05$, $\eta_G^2 = 0.07$		(ii) *Ease* $F_{1,11} = 19.98$, $p < .001$		(iii) *Comfort* $F_{1,11} = 23.6$, $p < .001$		(iv) *Confidence* $F_{1,11} = 14.95$, $p < .001$		(v) *Acceptance* $F_{1,11} = 15.46$, $p < .001$	
comparisons		*diff (s)*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*
AFTER	BEFORE	-0.32	$< .05^{*}$	0.74	$< .001^{***}$	0.75	$< .001^{***}$	0.51	$< .001^{***}$	0.57	$< .001^{***}$

(b) SESSION × MANIPULATION

		(i) *Ease* $F_{2,22} = 5.66$, $p < .01$		(ii) *Comfort* $F_{2,22} = 9.99$, $p < .001$		(iii) *Confidence* $F_{2,22} = 6.83$, $p < .01$		(iv) *Acceptance* $F_{2,22} = 7.11$, $p < .01$	
comparisons for FLIP		*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*
AFTER	BEFORE	1.54	$< .05^{*}$	1.67	$< .001^{***}$	1.29	0.29	1.25	$< .05^{*}$

A.3 Experiment 2: Results for Sitting versus Standing

Table 6: Main effect. Note: Only measures with significant difference are shown.

(a) MANIPULATION

		(i) *Time* $F_{1.09,11.96} = 34.49$, $p < .001$, $\eta_G^2 = 0.47$		(ii) *Smoothness* $F_{1.32,14.54} = 6.11$, $p < .05$, $\eta_G^2 = 0.12$		(iii) *Ease* $F_{2,22} = 81.44$, $p < .001$		(iv) *Comfort* $F_{2,22} = 83.34$, $p < .001$		(v) *Confidence* $F_{2,22} = 45.66$, $p < .001$		(vi) *Acceptance* $F_{2,22} = 42.03$, $p < .001$	
comparisons		*diff (s)*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*	*diff*	*p-value*
ROTATE	SPIN	-1.21	$< .001^{***}$	-0.007	$< .001^{***}$	1.41	$< .001^{***}$	1.62	$< .001^{***}$	0.93	$< .001^{***}$	0.83	$< .001^{***}$
ROTATE	FLIP	-1.03	$< .001^{***}$	-0.0048	$< .01^{**}$	1.73	$< .001^{***}$	1.98	$< .001^{***}$	1.73	$< .001^{***}$	1.64	$< .001^{***}$
SPIN	FLIP	0.18	1	0.0022	1	0.32	.12	0.36	.13	0.8	$< .01^{**}$	0.81	$< .01^{**}$

(b) DIRECTION

		(i) *Time* $F_{1,11} = 22.59$, $p < .001$, $\eta_G^2 = 0.14$	
comparisons		*diff (s)*	*p-value*
ADD	ABD	-0.47	$< .001^{***}$