Although free finger interaction seems to require a similar layer setup as previous work, the finger movement differs from the aforementioned techniques in some crucial aspects, therefore jeopardizing the applicability of their guidelines. While a minimum 4 cm thickness may apply for the bimanual usage of magic lenses while standing at tabletops [15], this thickness may not apply in a seated position, where only one hand is used. Although it is plausible that the thickness acquired by Kattinakere et al. [8] would apply to free finger movement, this assumption cannot be made, as a finger moves along a different trajectory than a stylus. The movement path of a finger is more curved than that of a stylus, making the path longer, and therefore more prone to drifts [4].

Additionally to the inherent differences of free finger interaction, several other factors differ from previous work. While Spindler's setup increases the stability of the object by spreading the load over two hands [16], the stability is potentially reduced by the user having to hold the object in midair without allowing arms to be rested. The continuous visual feedback provided in Kattinakere's study allows closed-loop adjustment of the height, potentially resulting in a more precise movement [8].

Another important factor of near-surface input is the engagement technique, analogous to clicking. Existing engagement techniques that have been proposed for midair input allow at most one input state per hand, and can be classified as follows. First, a hardware button can be used. This can be either on a separate device, e.g., Mysliwiec used one hand to point, the other to press a clicking-key [12], or be on the pointing device, e.g., Subramanian used a button on the stylus for clicking [17]. Second, the movement can be used to change state, like crossing in and out of targets (horizontal movement), or crossing midair layers (vertical movement) [17]. Finally, hand-shape gestures can be used to decouple the movement from the engagement. Wilson used a pinching gesture of index finger and thumb to click [19]. In order to improve the stability of the tracked cursor, Kato and Yanagihara tracked the knuckle position instead of the finger tip, making the pinching engagement independent from the movement [7]. Vogel used the striking of the thumb on the index finger to detect a click [18]. This was improved by Banerjee et al. to striking the middle finger instead of index finger to enhance stability of the cursor movement [1]. Perhaps the most intuitive of all engagement techniques is to emulate tapping by using the same finger motion in the air [18]. This was ranked best by users because of the familiarity with mouse clicking [3]. Pyryeskin et al. performed a gesture elicitation for above multitouch-surface selection. Air-tapping ("push with a finger") was the second most frequent gesture (26.6%), slightly less frequent than grabbing (35.2%), which occupies the whole hand [14]. Our paper takes a closer look at this index finger tapping, and addresses the shift of the finger that can cause erroneous input.

Before tackling the challenges of layer access and engagement, we will determine the required layer thickness for maintaining the tracking state in the next section.

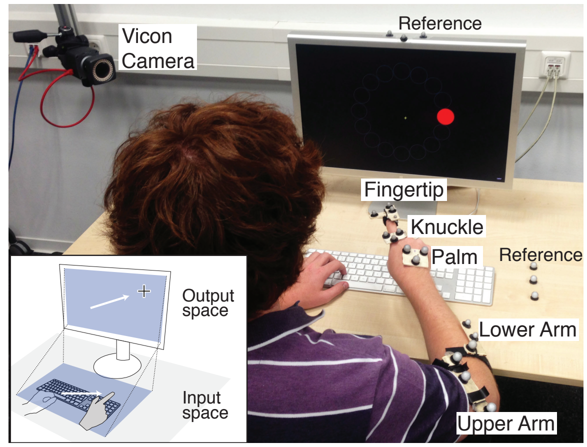

Figure 2: The physical setup of the user studies. Reflective markers were attached to the participants to record data for various sections of the right arm. Indirect mapping was used to prevent effect of hand occlusion.

STUDY 1: THICKNESS OF ABOVE-SURFACE LAYERS

In layered mid-air interactions, having thin layers allows for a large number of such layers within a given interaction volume. However, we need to take into consideration the ergonomics of the human arm and hand. Factors like hand tremor and drifting, and lack of haptic feedback in the near-surface space, make it difficult for users to maintain their hands at a constant level. This makes thinner layers harder for users to stay inside. This study aims to determine a suitable thickness of such near-surface layers while the finger remains in the tracking state (Fig. 1 transition 1).

Apparatus

The position of the users' index fingertip was used as input to the application, to control a screen cursor. To obtain accurate positional coordinates, we used a Vicon motion-capture system to track passive infrared-reflective markers, which provided three-dimensional data with sub-millimeter accuracy at 100 Hz. Markers were attached to the user's finger with lightweight patches (<8g each). Before the test, users flexed their fingers to ensure that the patches did not inhibit their movements. In addition, we recorded the position of the user's index fingertip, wrist and elbow for analysis purposes (Fig. 2). The position of knuckle and palm was only recorded in the second study. The experiment tasks were displayed on an Apple Cinema Display (49.5 cm × 30.5 cm; 1920 × 1200 pixels). To prevent the influence of hand occlusion, the user's finger was mapped from the orthogonal projection onto the desk surface, to the vertical screen. The participants sat in front of a desk, approximately 50 cm away from the screen, on an adjustable chair with armrests. The height of the chair was adjusted by each user according to preference.

Participants

Eight right-handed volunteers (two female) of computer science background were recruited (mean age of 24). Participants did not exhibit any known hand disabilities (severe tremor, etc.). Any visual impairments were noted down prior

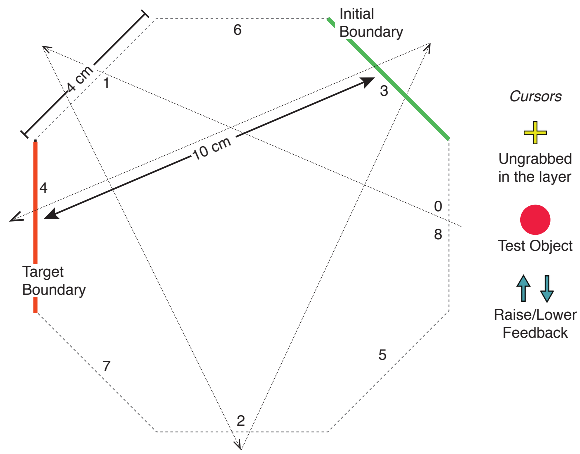

Figure 3: The steering tasks performed by the users. The cursor was used to grab a test object, and drag it from the initial to the target boundary. Different cursors provided feedback about the z-direction in which the finger needed to be moved to re-enter the layer.

to the test and it was confirmed that they had been corrected. All participants had a computer science background.

Task

The task users performed involved finger steering within a near-surface layer, similar to Kattinakere et al.'s [8]. Users controlled an on-screen cursor with their index finger to drag the object from an initial boundary across a target boundary (Fig. 3). After crossing each boundary, the next target boundary was shifted anti-clockwise. When the cursor enters the object, the object was automatically grabbed (no engaging gesture required). When the finger exited the layer, the object was released. Users could only pick the object back up again by readjusting their finger height.

Whenever the finger was outside the near-surface layer, the cursor changed to an arrow indicating whether the finger was too high or too low. Therefore, during the movement within the layer, the users needed to rely on proprioception to maintain the height of the finger.

To ensure constant initial velocity of hand movement, a short pause was programmed after crossing each goal. Since arm movements vary depending on the direction, the eight boundaries were evenly distributed, in order to make sure that the results were applicable to 2D movements in general.

Design

A within-subjects design was used for the user study. The independent variables used were the *Thickness*, *SurfaceSupport*, and *MovementDistance*. The different layer thickness values used were 1, 2, 3 and 4 cm. The two hand support configurations tested were (1) the users rested the wrist on the desk surface, and (2) the wrist had no support. Two types of movement sizes were used *SmallMovement* (1 cm) and *LargeMovement* (10 cm).

- *DriftCount*: Less drifts indicate that the user can reliably maintain the finger in the layer. This is the main measurement for determining optimal thickness. However, less drifts may results from the fact that the user carefully moves the finger during the test.
- *ManipulationTime*: Time taken for dragging the object from the initial boundary to the final boundary. This is measured only when the finger is within the given thickness and the object has been grabbed. This time excludes instances where the finger drifts outside the layer, causing the object to be released. Lower cursor manipulation time results from faster movement within the layer. This indicates that the user is more confident in moving the object in the layer.

In summary, the experimental design was: 8 participants $\times$ 4 *Thickness* $\times$ 2 *SurfaceSupport* $\times$ 2 *MovementDistance* $\times$ 8 boundaries $\times$ 3 repetitions = 3072 total trials.

Data analysis

In order to make each row of the data represent the performance across all movement directions, we averaged the *ManipulationTime* across all movement directions in the same repetition. *DriftCount* was also summed across movement directions.

We used mixed-effect model analysis of variance (henceforth "ANOVA") in which *Thickness*, *MovementDistance*, and *SurfaceSupport* are fixed effects and *UserID* is a random effect. We used Tukey HSD for post-hoc tests (henceforth "post-hoc test"). For all significant interaction effects, we performed the post-hoc tests for both main and interaction effects. The statistically significant results in the following section have the agreement of post-hoc tests up to the highest degree of significant interaction effects.

As expected, both *DriftCount* and *ManipulationTime* are not normally distributed. Appropriate transformations were applied before ANOVA and post-hoc tests. Neither Poisson nor Gamma-Poisson distribution fitted the *DriftCount*; therefore, it is aligned-rank transformed [21]. The *ManipulationTime* fits log-normal distributions, so we applied a $y = log(x)$ transformation.

Due to the lack of normality, descriptive statistics (mean, median, and 95% confidence interval) derived directly from the data may not be an accurate representation, e.g., means of *ManipulationTime* are pulled higher by a long tail of the log-normal distribution. The resolution of the statistics could be too low, e.g., the median of many *DriftCount* conditions are 1 regardless of their distributions. Also, simple inverse-transforms of CIs may produce the interval that excludes the sample mean. Therefore, to provide useful estimates, we calculated the descriptive statistics by ordinary non-parametric bootstrapping (10,000 replicates). CIs were calculated with the bias-corrected and accelerated method (BCa).

We used the original scales for the charts and the descriptive statistics. Error bars denote 95% confidence intervals.

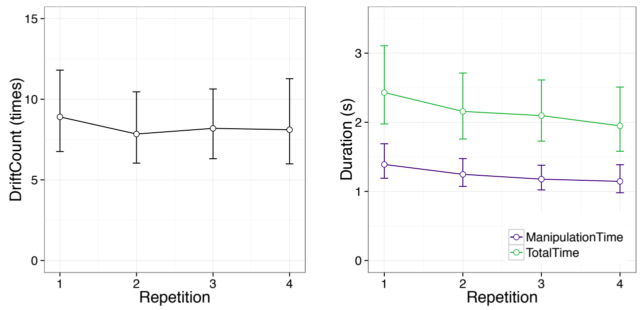

Figure 4: User's performance in the first repetition differs from the others, indicating the learning effect. The learning curve stabilizes after the first repetition.

Table 1: The significant main- and interaction effects to the both *DriftCount* and *ManipulationTime*.

	df		*F*	*p*
DriftCount				
Thickness	3	329	79.534	<.0001
Height	1	329	171.09	<.0001
MovementSize	1	329	162.01	<.0001
Thickness * Height	3	329	60.57	<.0001
MovementSize * Thickness	3	329	53.87	<.0001
MovementSize * Height	1	329	110.95	<.0001
MovementSize * Thickness * Height	3	329	28.88	<.0001
ManipulationTime				
Thickness	3	329	41.68	<.0001
Height	1	329	330.38	<.0001
MovementSize	1	329	215.10	<.0001
Thickness * Height	3	329	11.35	<.0001
MovementSize * Thickness	3	329	20.59	<.0001
MovementSize * Height	1	329	29.10	<.0001

Results and Discussion

Learning Effect

As expected, we found a statistically significant learning effect of repetitions on *ManipulationTime* ($F_{3,441} = 4.91, p = .0023$). The post-hoc test indicated that only the first repetition stood out from the rest (16% slower in *ManipulationTime*, Fig. 4). The faster movements suggest that the users were rapidly gaining confidence after one repetition whereas the drifts occurred independently of the users' confidence ($F_{3,441} = 1.12, p = .3394$). The interaction effects between repetitions and other independent variables were not statistically significant. To rule out the influence of users' confidence, we excluded the first repetition from the following analysis (resulting in 384 data rows).

Layer Thickness Analysis

The ANOVA results are shown in Table 1. As expected, we found significant main effects of *Thickness*, *SurfaceSupport*, and *MovementDistance* on both dependent variables. Due to significant interaction effects, we will take a closer look at each condition to determine suitable thickness.

Movement with surface support
When the hand was supported by the desk surface, post-hoc tests found no statistically significant differences in *DriftCount* and *ManipulationTime* among the 2, 3, and 4 cm thicknesses. The users drifted more in the 1 cm layer (M = 6.50 times) than in the others (4.46 times). This difference was independent of the movement sizes (Fig. 5). The users also moved 29% slower in the 1 cm layer than the others. Post-hoc comparison in the *SmallMovement* condition, which is the easiest condition, highlighted the worsened performance of the 1 cm compared to the 2 cm layer. When users achieved comparable *DriftCount*, the post-hoc test indicated that the *ManipulationTime* is significantly slower. In summary, when surface support is available, the near-surface layer can be as thin as 2 cm without worsening the user's performance.

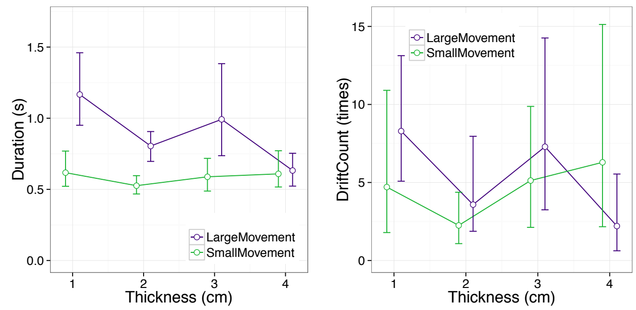

Figure 5: Drift occurrences and manipulation time for hand-on-desk conditions. Post-hoc tests shows that the 2 cm thickness yields similar performance to thicker layers.

In contrast to [8], our visual feedback only indicates whether the finger is in the layer, too high, or too low. The similar result of the 2 cm thickness suggests that maintaining a finger in these conditions may not require continuous visual feedback of the finger height with respect to the layer.

Movement without surface support
When the hand and arm are not supported by the desk surface, the effect of *Thickness* was different between *SmallMovement* and *LargeMovement*.

SmallMovement: Post-hoc tests found no statistically significant differences in *DriftCount* and *ManipulationTime* among the 2, 3, and 4 cm thicknesses. Fig. 6 shows that drifts in the 1 cm thickness (M_{1cm} = 10.79 times) triples those of the others ($M_{2,3,4cm}$ = 3.25 times). Longer *ManipulationTime* in the 1cm layer (M_{1cm} = 1.17 s vs. $M_{2,3,4cm}$ = 0.95 s) suggests that users moved with significantly less confidence for this thickness.

Although the 2 cm result is similar to the hand-on-surface condition, the *SmallMovement* in midair is only comparable to the *LargeMovement* with surface support ($M_{DriftCount}$ = 4.33 times and $M_{ManipulationTime}$ = 0.81 s). *SmallMovement* with surface support in the 2 cm thickness is slightly better than both ($M_{DriftCount}$ = 2.25 and $M_{ManipulationTime}$ = 0.53).

LargeMovement: Post-hoc tests found no statistically significant difference in *DriftCount* between the 3 cm and 4 cm thickness (M_{3cm} = 9.58 times, M_{4cm} = 4.15 times, Fig. 7). However, users were faster in the 4 cm (M_{3cm} = 1.91 s, M_{4cm} = 1.10 s). The 4 cm thickness in midair also performed simi-