Investigating Eyes-Free Recognition and Distinguishability of Textile Icons in Pairs

Oliver Nowak
RWTH Aachen University
Aachen, Germany
nowak@cs.rwth-aachen.de

René Schäfer
RWTH Aachen University
Aachen, Germany
rschaefer@cs.rwth-aachen.de

Elisabeth Jane Buttkus
RWTH Aachen University
Aachen, Germany
elisabeth.jane.buttkus@rwth-aachen.de

Lea Emilia Schirp
RWTH Aachen University
Aachen, Germany
lea.schirp@rwth-aachen.de

Heiko Müller
University of Oldenburg
Oldenburg, Germany
hilfeheiko@gmail.com

Jan Borchers
RWTH Aachen University
Aachen, Germany
borchers@cs.rwth-aachen.de

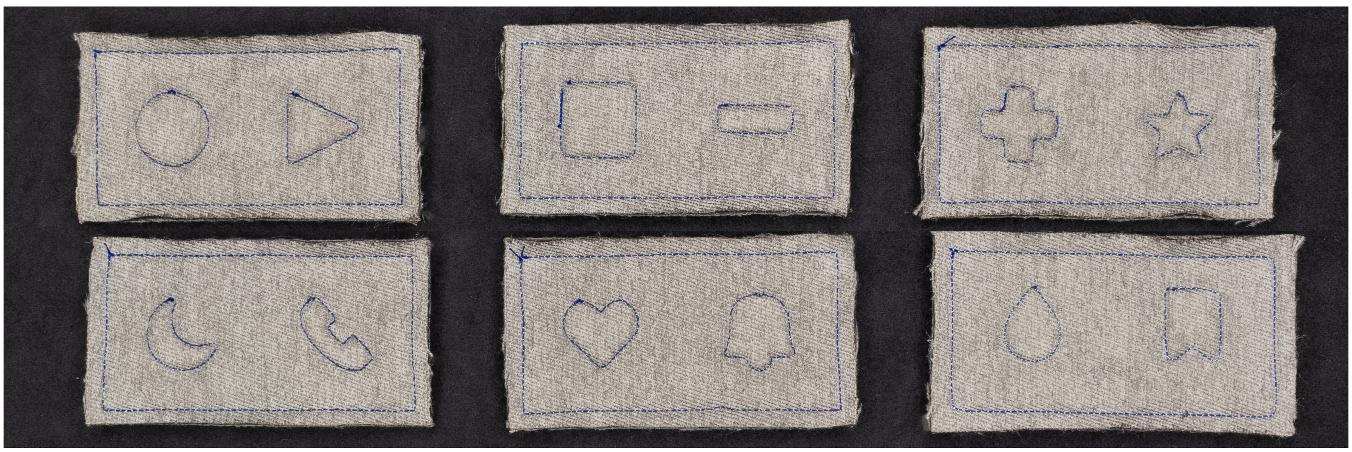

Figure 1: Examples of the many textile icon pairings from our study. They include all icons evaluated: *Circle*, *Triangle*, *Square*, *Minus*, *Plus*, *Star*, *Moon*, *Phone*, *Heart*, and *Bell*. *Raindrop* and *Bookmark* were only used in the familiarization phase of the study. Each combination of the icon shapes was tested in our study. Icons were fabricated according to [31].

Abstract

Textile interfaces enable silent and discreet input on clothing, accessories, and smart home furniture. While researchers already presented approaches to make them technologically feasible, it is not fully clear how users experience textile interfaces and how well users perform when vision-free usage is encouraged. Recently, designs of single textile icons, i.e., symbols used as textile buttons or labels, were investigated. Practical user interfaces, however, typically consist of entire groups of nearby icons. Their haptic distinguishability is key for seamless operation. Furthermore, it is unclear whether icon recognition benefits or suffers when comparing neighboring icons is possible. We conducted a study where users blindly palpated icon pairs, tried to recognize the individual shapes and rated how easy they were to tell apart. We present our observations on haptic distinguishability, which, inter alia, show that more haptic cues via neighboring icons do not impact shape recognition.

TEI '25, March 04–07, 2025, Bordeaux / Talence, France

ACM ISBN 979-8-4007-1197-8/25/03
https://doi.org/10.1145/3689050.3704931

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; **User studies**; **Haptic devices**.

Keywords

Textile Interfaces, Icons, Eyes-free Interaction, Haptic Recognition

ACM Reference Format:
Oliver Nowak, René Schäfer, Elisabeth Jane Buttkus, Lea Emilia Schirp, Heiko Müller, and Jan Borchers. 2025. Investigating Eyes-Free Recognition and Distinguishability of Textile Icons in Pairs. In *Nineteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '25), March 04–07, 2025, Bordeaux / Talence, France.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3689050.3704931

1 Introduction

Textile interfaces are well suited for controlling smart homes as they can be embedded into many everyday objects in the home, such as pillows, sofas, or table runners. They can provide versatile shortcuts, e.g., for communication and media playback or control smart home devices such as lights or curtains. In the home, they offer a silent, in-place, and more discreet alternative to voice and traditional remote controls. While there is a solid body of research on how to fabricate such functional textile interfaces for input [1, 5, 25, 26, 28, 33] and output [9, 22], empirical evidence to inform design guidelines for such interfaces is sparse. Especially the haptics of textiles can be used to enable the eyes-free use of the interface and should help communicate what actions users can perform. This is essential to make textile controls successful in real-world scenarios. Mlakar and Haller [20] and Schäfer et al. [31] previously investigated such design guidelines for the recognizability of *textile icons*—i.e., icons and shapes known from classical GUIs but implemented on textile surfaces—since they provide quick-to-understand and space-efficient signifiers for textile interfaces. While low recognition rates were observed for flatly embroidered icons [20], satisfactory recognition rates were found for icons standing out from the fabric [31]. Both discuss shape confusions, i.e., haptic shapes being recognized as another symbol; however, targets in their studies were palpated and recognized only individually, one at a time. This, however, is different from how users will come upon such elements in real-world scenarios. There, multiple textile UI elements will be placed close to each other, and it is unclear whether participants will encounter synergies between neighboring icons. Furthermore, from studies with isolated icons, it remains unclear whether the original icon and a participant's confused answer actually feel different—more precisely, whether designers could use those icons next to each other without making people struggle to differentiate them. We expect people to better understand the features of the element in focus due to the comparison with the features of surrounding UI elements. If this is the case, interface designers could use this knowledge to arrange icons in their interfaces to support feature recognition without substantially changing the individual icon design.

We conducted a user study in which participants recognized textile icon pairs without looking and rated their haptic similarity to understand what makes icons feel different and whether people use information from one icon to understand another better. Overall, we hypothesize the following:

(1) Simple geometric shapes such as *Square*, *Circle*, and *Triangle* will overall perform better due to their simplicity.
(2) Users will better recognize icons in pairs since they will compare shape features of both icons with each other.
(3) Recognition success of icons will vary depending on their neighbors.
(4) If our icon pairs contain the same icons twice, users will identify them faster.
(5) If icons look mostly the same and only differ in the presence of a few additional features (*Moon-Phone*, *Triangle-Heart*), those features will stand out, leading to good distinguishability and fast recognition times.
(6) If two icons differ in volume, this will help users understand the icon sizes and improve recognition performance.

2 Related Work

Current research on textile interfaces spans a wide variety of topics, from clothes [8, 25, 33] and car interiors [11] to surface gestures [19], manufacturing and prototyping techniques [6, 12, 28], knitting [18], weaving [29], stateful interfaces [21], and more. That research, however, has primarily focused on new textile artifacts and their technical underpinnings. Deriving design guidelines has received little attention so far.

2.1 Haptic Shape Recognition

Physical shapes can use several haptic properties to distinguish UI elements from one another. Such separating cues use different intensities of salient features. Lederman and Jones [16] surveyed tactile illusions, including categories that affect the perceived *texture* and *size* of shapes. For example, the *Müller-Lyer haptic illusion*, due to which two lines are perceived as differing in length due to their endpoints, also applies to touch. For 2D outlined shapes, Larsson et al. [13] let people visually perceive 44 unknown closed shapes and name what they considered the main part of each shape. They found that, on a visual level, small inconsistencies in shapes, such as irregular bumps, are often considered salient.

Haptic shape recognition is often investigated using raised-line drawings [e.g., 3, 10, 15]. In Lebaz et al. [15]'s investigations, they found 46% of participants successfully identified shapes like keys and saws with an average response time of 86 s. This is in line with the findings of Kalia and Sinha [10], who found the complexity of a shape influences its recognition, as it creates more cognitive load to memorize local details and combine them into a model of the whole shape. Ng and Chan [23] investigated the ability of people to discriminate between different haptic geometric shapes with a height of 4 mm. They found simple shapes (e.g., *circle* and *square*) and shapes that had a smaller number of edges to be perceived significantly faster than complex shapes like a *six-pointed star*. Plaisier et al. [27] found that for 3D shapes, the presence of edges and vertices is an important salient shape feature. This makes *height* also attractive for shape recognition on textile surfaces as it creates these salient features for shapes relative to their surroundings. Using height as a key distinction was also investigated by Leo et al. [17] with pin-array displays. In their study, sighted people needed significantly longer than blind people to respond to a new shape (47 vs. 17 seconds).

Mlakar and Haller [20] state that elements that differ in *height*, *shape*, or *texture* will stand out compared to their surroundings and other controls. They identified *height* as the easiest way to create haptic contrast. Additionally, the authors suggest designing shapes to be as simple as possible, which aligns with other previous [4] and more recent [31] work. Using only an outline of a shape without further haptic cues is not sufficient for reliable recognition [20, 24, 31], especially if recognition has to happen quickly, as with frequently used everyday interfaces.

2.2 Icons

Graphical icons should resemble real objects to become universal and intuitive for their user group [2]. For textiles, Holleis et al. [8] started investigating different types of symbols on an apron. They investigated *visible*, *ornamental*, and nearly *invisible* buttons in a study in which participants controlled home cinema devices. Participants liked the subtlety of invisible buttons and that they did not need to be in a specific location near the target device to use it. However, when using them blindly, participants considered them barely usable. For pneumatic buttons, Harrison and Hudson [7] found that flat buttons needed more visual attention than their raised, recessed, and classic counterparts, thus making them harder to use. Mlakar and Haller [20] support these findings with results from a study in which they asked people to recognize four different symbols stitched as outlines on a piece of fabric. Half of their 30 participants could recognize a star, and 11 understood a heart, while the other shapes were barely recognized correctly. Their findings suggest that only using a *flat* version of an outlined shape is insufficient for reliable recognition. Furthermore, they found that shapes should be at least 13 mm in size and stand out 1.6 mm above the surroundings. Schäfer et al. [31] built upon that work investigating *textile icons* that consisted of laser-cut templates enclosed in fabric, resulting in three different height profiles (*flat*, *raised*, and *recessed*). They found that people were faster in recognizing shapes raised above their surroundings, and their ratings showed a tendency for filled shapes compared to outlined shapes.

3 Study

In our user study, we investigate the effects of neighboring textile icons on recognition time and success. For this, we investigate ten textile icons adopted from [31]. We replicated their raised icons since they were already investigated individually, performed well, and were subtle and simple, which are criteria drawn from [4, 8, 20]. We selected the shapes based on their performance and their applicability in smart home interfaces: ***Minus, Plus, Star,*** and ***Heart*** could be used, e.g., to control volume and favor media. The other icons could be used in communication applications to call someone (***Phone***), manage notifications (***Bell***), and set a do-not-disturb mode (***Moon***). ***Circle, Triangle,*** and ***Square*** represent the simplest and probably most easily recognized shapes. ***Raindrop*** and ***Bookmark*** are also adopted from [31] but were only used by participants to familiarize themselves with the study procedure. We removed *Lightning* and *Arrow* due to their low recognition and to keep the study duration reasonable for our participants. Figure 1 shows all shapes used. In practical user textile interfaces, those icons could be combined in two ways: Icons could create semantic pairs for controlling the state of a device. For example, *Plus*/*Minus* could be used for changing values, or *Triangle*/*Circle*/*Square* could be used for navigation like on Android phones. However, they also could be used for device and application selection before manipulation takes place. Then, rather unrelated icons like *Phone* and *Bell* could be placed next to each other. Since good distinguishability is necessary in any case, we decided not to constrain our icon combinations by testing all possible combinations. This allows us to compare icons with very different shape features as usually, semantic icon pairs have a similar body style (e.g., both *Plus* and *Minus* consist of only straight lines).

We fabricated our icon pairs following the process in [31]: We laser-cut the shapes from medium-density fiberboard and glued them onto fabric with the shape outlines already embroidered onto it. Then, we placed another fabric on top and embroidered the

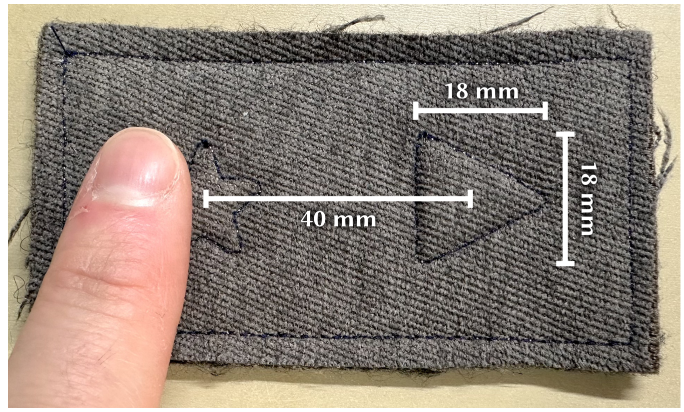

Figure 2: Dimensions of an exemplary icon pair. Each icon filled a square of 18 mm×18 mm. The centers of the icons were 40 mm apart. A finger palpates the icons for a size comparison.

outlines again with an offset of 1 *mm*, which created clear edges. We used straight stitches with thin yarn for all embroidery. We chose furniture fabrics for both layers following [24, 31]: a 100% polyester with a fine texture and a weight of 270 g/m^2 that should avoid friction burn. We used a Bernina 880 automated embroidery machine. Each icon had a height of 1.6 mm above the base fabric. Each shape's longest dimension was 18 mm to make all icons a similar size. Within each pair, icons were placed at a distance of 40 mm between their centers, creating a space of approximately 2 cm between the icons. We decided on this distance as it clearly haptically suggests that those elements do not form one unit but still feels close and quick to reach over. Figure 1 shows six icon pairs containing all ten icons from our study and the two used for familiarization. Figure 2 shows an examplary icon pair and its dimensions.

3.1 Study Setup and Apparatus

Our study setup (Figure 3) consisted of a sight protection wall and an interaction area. The interaction area was always placed to the side of the hand the participants wanted to use for the recognition. The participants sat on the other side of the sight protection and reached over to the interaction area so they could not see the icons by accident. The interaction area was covered with padding foam with a button for time measurements embedded close to the sight protection and some Velcro to attach the icon pairs. This button was connected to an Arduino Uno, which sent all raw data to a connected laptop via a serial connection. A camera recorded the participants' hands.

3.2 Study Procedure

In the beginning, we introduced each participant to the context of textile icons and interfaces, i.e., their usage in smart homes. After signing an informed consent form, they filled out a demographics questionnaire. To familiarize the participants with the haptic feeling of the icons and the study setup, they explored a sample icon pair

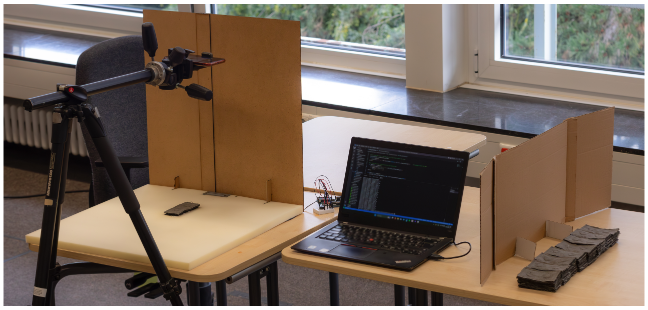

Figure 3: Study setup. Participants were seated on the chair to the left. They interacted with an icon pair, which was attached to an area covered with padding foam. The large visual barrier prevented participants from looking at the icon pair. The experimenter was seated at the laptop and with an icon collection behind the small visual barrier on the right. To measure interaction times, participants pressed a large 3D-printed button (grey, near the large barrier) connected via an Arduino to the laptop.

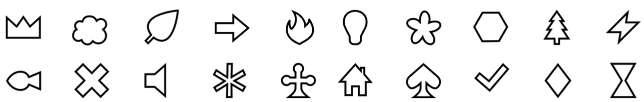

Figure 4: The 20 unused icons from the questionnaire that was given to participants after the trial phase. Top, left to right: *Crown, Cloud, Leaf*, *Arrow, Fire, Bulb, Flower*, *Hexagon, Tree, Lightning*. Bottom, left to right: *Fish, Cross, Speaker, Asterisk, Clubs, House, Spades, Check, Diamond, Hourglass.*

containing *Raindrop* and *Bookmark*. These shapes were not part of the study.

For each trial, the experimenter placed an icon pair on the Velcro. Participants pressed a button to start and finish the time measurement for a trial. They were instructed to press the button to finish a trial once they could name or at least describe both icon shapes after palpating them. Then, they described both shapes to the experimenter as accurately as possible. Participants did not receive any feedback on their recognition and were asked to only use their preferred hand for such tasks to ensure consistent results. This process was repeated for each of our 55 icon pairs with a planned break after 28 trials. Unlike in [31], participants were unaware of the possible shapes they would encounter during the study to ensure that our similarity measurements were purely based on haptic similarities.

After all trials, participants received a shuffled list of 30 different shapes consisting of shapes from the study and those pictured in Figure 4. They marked all shapes they believed had been used during the trials. Participants were not informed how many different shapes occurred during the study to avoid them adding or excluding shapes based on this information. Afterward, they filled out a questionnaire for all shapes they thought they had encountered.

3.3 Variables

We controlled the icon pairs that participants encountered during the study. The participants palpated 55 icon pairs (all possible icon combinations without mirrored pairs) in random order. We measured the *recognition time* using the embedded button, their reported *icon recognitions*, *misidentifications* with other shapes, and participants' reported *icon distinguishability*. At the end, we collected *questionnaire data*. If a trial took over 60 seconds, we reported this as *time out* and set their recognition time to 60 seconds.

3.4 Participants

Overall, 28 people participated in our study (M=26.5 years, SD=5.8 years. 13 self-reported as male, 12 as female, 2 as others, and 1 as n/a). 27 participants were right-handed, and 1 was left-handed. 20 participants were computer science students of different focuses. Furthermore, 2 engineering, 1 biology, and 1 elementary school education student participated. 1 person reported as housewife and sewing apprentice, 2 as researchers (no specialization specified), and one person did not provide this information. The participants could provide their answers in their native language or English. All participants volunteered without monetary compensation.

3.5 Results

Since we were interested in both the influence of individual icons in one pair and the overall performance of icon pairs, we analyzed task time and recognition rates for single icons and icon pairs. For individual icons, we calculated the average recognition time of all *icon pairs* that included the corresponding item. For simplicity, we call this measurement *icon time influence*. For icon pair recognition rates, we calculated the average recognition success of the icons contained, with 1 for successful recognitions and 0 otherwise.

Since our measured recognition times were log-normally distributed, we analyzed the log-transformed data using repeated measures ANOVA and paired t-tests with a Holm correction as post-hoc tests. All other measurements were analyzed using Friedman tests and Wilcoxon signed-rank tests with Holm corrections for the post-hoc analysis [14, 30].

3.5.1 Recognition times. Figure 5 shows the average icon pair *recognition time*. The ANOVA test revealed significant effects of icon pairs on *recognition time* ($F(54, 1458) = 3.441$, $p<0.001$). All significant differences involve the three worst-performing icon pairs: *Bell-Bell*, *Star-Bell*, and *Plus-Bell*. Those performed significantly differently from most of the ten best icon pairs. In the appendix, we included a clearer heatmap representation in Figure 10.

For *icon time influence*, the ANOVA test revealed significant effects ($F(9, 243) = 12.16$, $p<0.001$). The post-hoc test mainly revealed that *Square*, *Minus*, and *Moon* led to a significant better *icon time influence* than *Heart*, *Plus*, *Phone*, and *Bell*. All results are illustrated in Figure 6 (right).

3.5.2 Icon recognition. Although participants were informed about our motivation for finding symbols for smart home controls, their recognized symbols only occasionally matched this domain and, thus, varied strongly per subject. Therefore, we decided to classify the participants' recognition as follows:

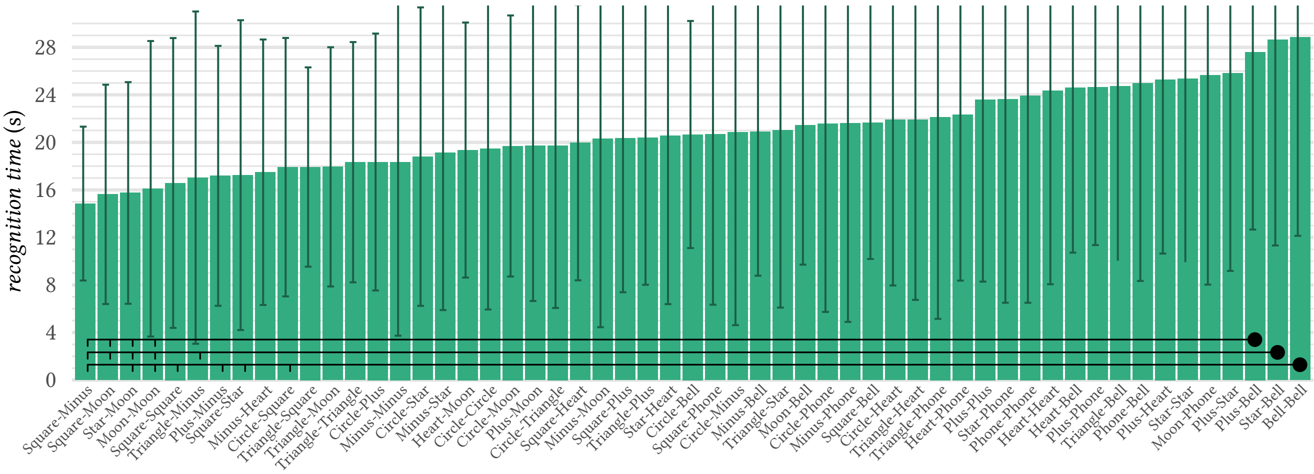

Figure 5: Average recognition times for our 55 icon pairs. Horizontal bars denote significant differences where all icon pairs connected with a tick are significantly different from the pair connected with a circle. Vertical bars are standard deviations. Overall, our icons performed similarly. The few significant effects only occurred when *Bell* was involved.

- *Recognition by name*: A response that contains a correct shape name but no other shape from our set. If participants additionally named incorrect details (e.g., a *Star* as "star with four spikes"), this was still accepted since the general shape was still interpreted correctly.
- *Recognition by description*: A description of the shape that describes all relevant features of the shape. This includes shape names such as "banana" and descriptions like "curved semi-circle" for *Moon*.
- *Incorrect recognition*: All shapes were counted as incorrect that were named and described incorrectly or whose description was too ambiguous (e.g., "Pac-Man ghost" for *Bell*).

We accepted alternative correct shape names for some shapes if the name's shape is an unambiguous shape description or an abbreviation. Therefore, we accepted "full moon", "oval", "ellipse" for *Circle*, "rectangle", "stroke", "ruler", "line", "bar", "dash" or "stripe" for *Minus*, "call button" for *Phone*, "cross" for *Plus*, "trapezoid" or "rectangle" for *Square*, and "play symbol/button" or "triangular arrow" for *Triangle*. If "rectangle" was used for *Minus*, we ensured that it was not used for *Square*. On two occasions, participants drew the shapes correctly instead of naming them; we accepted this as *recognition by description*.

The Friedman test revealed significant differences between icons on both the average *recognition by name* ($\chi^2(9)$=148.46, p<0.001) and *incorrect recognition* ($\chi^2(9)$=132.44, p<0.001). Figure 6 (left) shows how frequently our participants recognized the individual items and the significant differences for *recognition by name* and *incorrect recognition*. It is noticeable that, again, mainly *Bell*, *Phone*, and *Plus* were significantly harder to recognize than most of the other shapes.

3.5.3 Icon pair recognition. For icon pair recognition, we found significant effects between icon pairs ($\chi^2(54)$=736.78, p<0.001). However, except for two effects involving *Phone-Phone*, the significant effects all involve icon pairs with a *Bell* shape, which also performed worst overall. Figure 7 shows a heatmap of the total and individual frequencies of *recognition by name* for each icon pair. Particularly noteworthy is that the individual icon recognition count is approximately the same regardless of the tested icon pair. In the appendix, Figure 12 shows a detailed graph with all significant effects.

3.5.4 Shape misidentifications. To discuss misidentifications, we considered all recognition responses that clearly showed an association with existing shapes or objects. For this, we also considered objects where small features were added or removed (e.g., "egg with an edge at the bottom") and excluded path descriptions for this metric. We discarded shape transformations, like shearing or curving, since we found people to use the named shape as path description.

We found several shape misidentifications, the most common of which are listed in Table 1. It includes every misidentification that happened 11 times or more and, thus, must have been named by more than one participant. Similarly, to [31], misidentifications within our item set were directional: While *Plus* was confused as *Star* 130 times, the opposite direction occurred only 4 times. Similarly, *Phone* was confused as *Moon* 25 times, while the confusion in the opposite direction happened only once.

3.5.5 Icon distinguishability. We also found significant effects between icon pairs on *icon distinguishability* ($\chi^2(54)$=1010.9, p<0.001). Figure 8 shows the average ratings and significant effects. As expected, most of the twin pairs performed significantly differently from all other pairs. *Moon-Phone* and *Plus-Star* were the only non-twin pairs that were involved in a similar amount of significant effects. The only other significant differences between non-twin

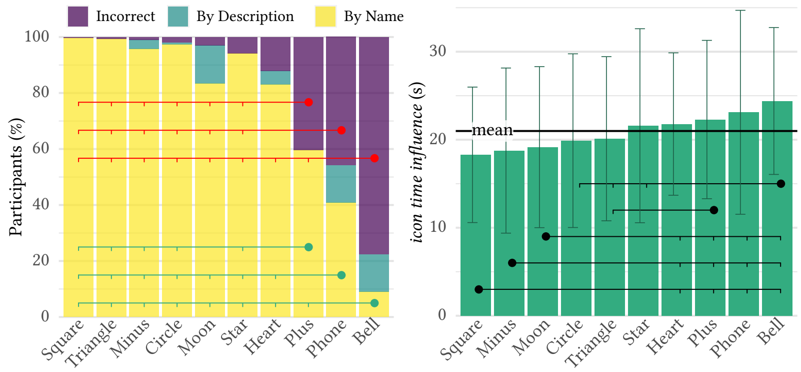

Figure 6: Icon-specific performance measurements: The left graph shows the relative frequency of how often a participant's response was classified as named correctly, described correctly, or incorrect. The right diagram shows the average recognition time of icon pairs including the corresponding icon. Horizontal bars denote significant differences. Here, all icons connected with a tick are significantly different to the icon connected with a circle. For icon recognition, green and red bars refer to significant effects for correctly and incorrectly identified icons, respectively. Vertical bars denote standard deviations.

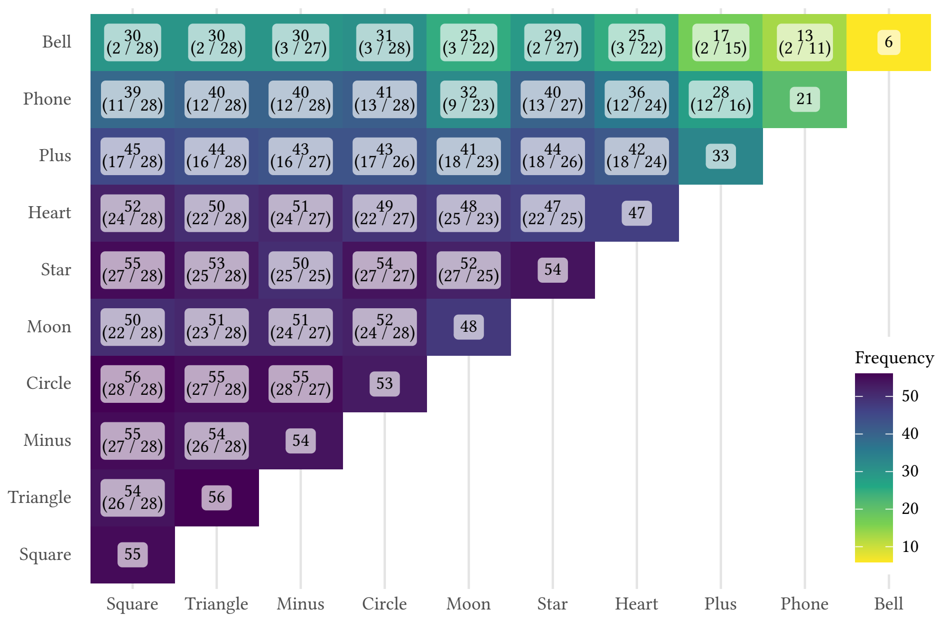

Figure 7: Heatmap showing the icon *recognition by name* per icon pair. The cells include the overall number of recognitions, as well as the number of recognitions for the individual icons in brackets in the form (row/column). One can see that icon pairs including *Bell* and *Phone* perform worse. Overall, individual icon performance stays approximately the same independent of the neighboring icon.

pairs were between *Star-Bell* and *Circle-Minus* & *Square-Moon*, between *Heart-Bell* and *Circle-Star* & *Square-Phone*, and between *Circle-Heart* and *Circle-Star*. Figure 11 in the appendix offers a clearer but less informative heatmap representation of the data.

3.5.6 Questionnaire Results.

Mapping Icons to Visuals. After the recognition task, we gave our participants a list of 30 different drawn symbols. From those, they had to mark the ones they believed to have appeared during the trials. *Square*, *Triangle*, *Moon*, and *Minus* were marked by all participants. Those were followed by *Heart* & *Circle* (M=93%, SD= 26%), *Star* (M=89%, SD=31%), *Plus* (M=82%, SD=39%), *Phone* (M=75%, SD=44%), and *Bell* (M=50%, SD=51%). Regarding the symbols not included in our icon set, 21% of our participants marked a *Cross* (SD=42%), 18% a *Flower* (SD=39%), and 11% an *Asterisk* (SD=31%). Other symbols were marked in 10% or less of all cases. Overall, 18 participants provided markings, which included all icons from our icon set.

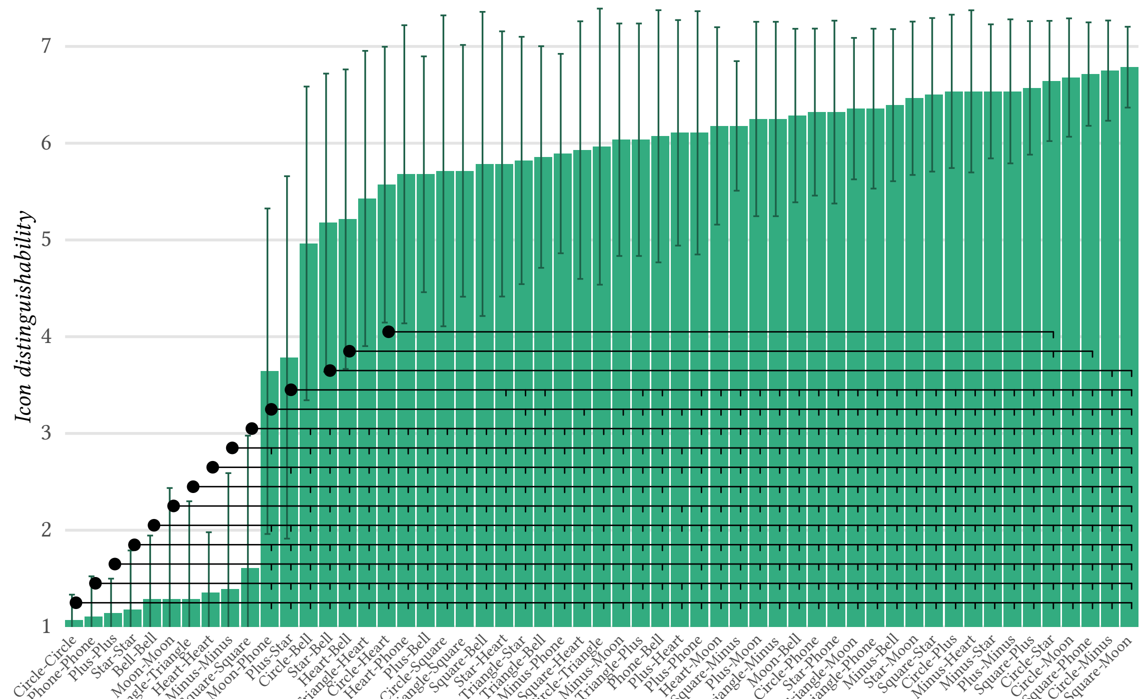

Figure 8: Ratings for *icon distinguishability* on a 7-point Likert scale ranging from 1 (very hard to tell apart) to 7 (very easy to tell apart). Vertical bars denote standard deviations. Horizontal bars denote significant differences, with all icon pairs connected with a tick being significantly different to the pair connected with a circle. While "twin pairs" were expected to be hard to tell apart, *Moon-Phone* and *Plus-Star* also received notably low distinguishability ratings.

Shape	Misidentified as
Bell	Ghost (83), Octopus (45), Circle with additional features (13), Semicircle with additional features (13), Fireball (11), Meteor (11), Upside-Down Tulip (11), Egg with additional features (11), Popsicle (11)
Plus	**Star** (130), Flower (13)
Phone	C (28), **Moon** (25), Semicircle (13), C with additional features (13), Sickle (12), Hook (11), Horseshoe (11), Nose ring (11)
Heart	V (22)
Moon	C (15)

Table 1: The most common misidentifications and how often they occurred, sorted in descending order of total misidentifications. Highlighted shapes were also included in our data set. Misidentifications that happened less than 11 times were filtered out. *Bell* leads this list with a total of 267 misidentifications (including those not listed).

Questionnaire Ratings. Our participants answered icon-independent and icon-specific questions if the icons were marked in the previous step, so their answers also belonged to the correct shape. Except for the *ease to recognize icons*, Figure 9 shows the results of those questions. To obtain a balanced and reliable test set, we excluded ten participants who had not marked all icons of our icon set on the questionnaire before from the significance analysis of the icon-specific questions. The Friedman tests revealed significant differences of the icons on *ease to recognize icons* ($\chi^2(9)$=60.018, p<0.001) and *confidence for correct icon recognition* ($\chi^2(9)$=63.49, p<0.001). The post-hoc tests showed only significant differences between *Bell* and *Triangle*, *Circle*, *Square* & *Minus* for *ease to recognize icons* and *Bell* and *Square* & *Minus* for *confidence for correct icon recognition.* Overall, our participants reported being confident when recognizing and telling icons apart. Unfortunately, the data shows no dominant exploration strategy for the task.

4 Discussion

Interestingly, our findings mostly contradict the hypotheses we specified initially. While *Plus*, *Phone*, and *Bell* aggravated recognition and caused the most significant effects, the small number of significant effects besides those icons demonstrate the complexity of designing for haptic recognition. In the following, we describe our observations for single icons as well as the effect of neighboring icons.

4.1 Single Icons

Simple geometric shapes do not outperform semantic shapes. On the one hand, the significant effects of *Plus*, *Phone*, and *Bell* on

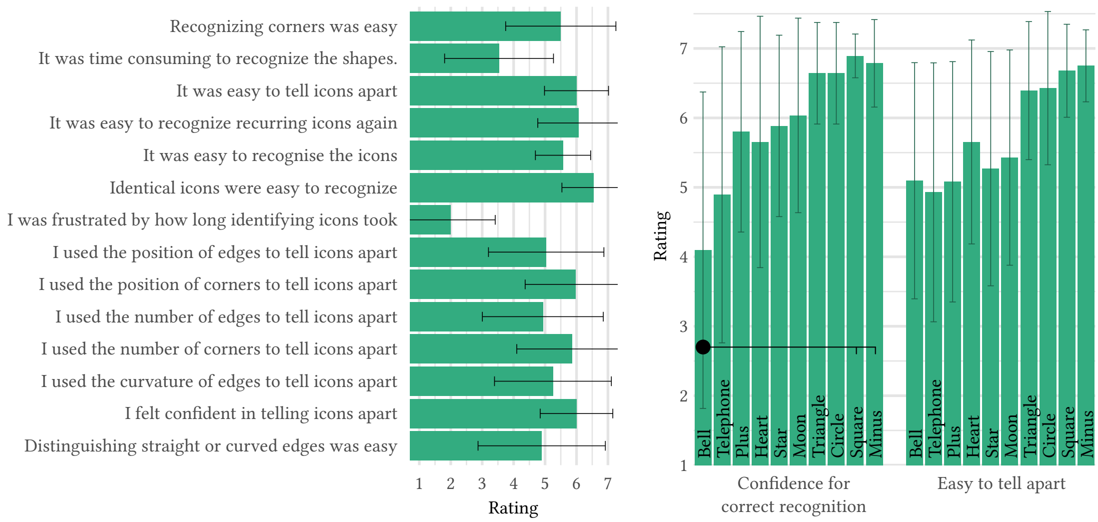

Figure 9: Average results of our icon-independent (left) and icon-dependent (right) questions on 7-point Likert scales. High rankings show high agreement and vice versa. Error bars denote standard deviations. Horizontal bars on the right denote significant differences, with all icons connected with a tick being significantly different from the icon connected with a circle. Overall, our icons were perceived as being easy to tell apart and to recognize without many significant differences between the icons.

recognition rates and *icon time influence* are in favor of Hypothesis 1. On the other hand, we found similar recognition rates for the more complex shapes *Moon*, *Heart*, and *Star* and our simple shapes *Square*, *Circle*, *Triangle*, and *Minus*. Considering *icon time influence*, *Circle* and *Triangle* do not show enough significant effects to allow a clear separation between simple geometric shapes and semantic ones. For *recognition time* and *icon pair recognition*, we also found no clear trends. Thus, although *Plus*, *Phone*, and *Bell* clearly show worse overall performance, our measurements for the remaining icons do not consistently show that Hypothesis 1 is true.

Users may understand complex shapes if they can imagine the application context more clearly. While we found that *Bell*, *Plus*, and *Phone* were recognized significantly less often, other measurements like their measured and rated distinguishability and participants' *confidence for correct icon recognition* show that it created a unique haptic sensation. Considering that Schäfer et al. [31] observed success rates of at least 73% for those shapes when users know them, and research for raised line drawings showed the importance of being able to visualize palpated shapes [3, 32], we hypothesize that such complex shapes will perform much better if users have a clear meaning or action associated with the icons.

Although subtle features are generally noticeable, misidentifications happen due to blending features. Similarly to [31], we found that confusions for *Plus*/*Star* and *Phone*/*Moon* solely occurred unidirectionally. Especially misidentifications from *Plus* to *Star* occurred noticeably often. Those confusions seem to occur since smaller features, like closely located corners, haptically "merge" into the overall shape. However, we could also observe that other subtle features were recognized. For *Bell*, for example, responses like "Fireball", "Meteor", or "Tulip" hint that our participants recognized even the subtly curved corners of our design. Those results match the results of [10] for raised line drawings, which stated that recognizability correlates with how well the shape can tolerate degradation in many cases.

The same icons could be used for different purposes. We found a high number of misidentifications for shapes that were associated with other objects. For *Bell*, this happened particularly often. We believe that due to its more filigree corners, the outline is harder to describe using simple paths or geometrical shapes, even visually, which forced our participants to find more abstract descriptions like "ghost", "octopus", and "fireball". Such icons could offer an opportunity to create haptic illusions and, therefore, be used manifoldly in different contexts for different meanings. Considering the *icon distinguishability* ratings, which showed acceptable distinguishability even for *Bell*, it could be used, for example, to announce emails or to start a gaming setup without messing with the mental image of the user, as long the user knows that the button is once located in the communication and once in the entertainment control section of the interface.

4.2 Icon Pairs

Significant time effects were present only when worst-case scenarios came together. We only found significant *recognition time* effects when *Plus-Bell*, *Star-Bell*, and *Bell-Bell* were involved. While our *icon influence* results show a significant negative influence of *Bell*, most *Bell*-pairs did not perform significantly differently from other pairs. Furthermore, we assume that the recognition of *Plus-Bell* and *Star-Bell* took significantly longer compared to, e.g., *Phone-Bell* since they require participants to palpate the complete outline for confident recognitions. However, this aspect alone did not lead to significant effects since already *Plus-Star* is not involved in significant effects. Therefore, it seems that the combination of bad recognizability of *Bell* and the necessity of palpating the whole outline explain this bad performance.

Identifying the same icon twice is not faster than recognizing two different ones. Our data contradicts Hypothesis 4 that states that twin pairs will be recognized faster for both simple and complex shapes because features would be re-recognized faster. Instead, such twin pairs distribute all over the performance range. Even more, while we expected the twin pair of an icon to be faster than pairs with other icons, shapes like *Bell-Bell*, *Star-Star*, and *Plus-Plus* show that icon equality does not improve recognition times.

Simple shape changes do not change how easy shapes are to tell apart. We expected two icons of the same overall shape with few simple changes to improve *icon distinguishability* and *recognition time* as participants would quickly understand the rough shape and the amount of differences is low (Hypothesis 5). Similar results were found for visual shapes where small inconsistencies were considered salient [13]. However, in our haptic task, *Triangle-Heart* and *Moon-Phone* received relatively low distinguishability scores from which *Moon-Phone* was in many cases significantly different. Also, regarding the *recognition time*, both did not perform significantly differently from most other pairs and belonged to the rather long-taking icon pairs. Therefore, from our collected data, Hypothesis 5 is incorrect.

Volume and extreme feature differences help but do not guarantee distinguishability. Looking at the significant effects on the *icon distinguishability* rating beyond twin pairs, *Moon-Phone*, and *Plus-Star*, we observed possible patterns for what makes icons feel different: *Star-Bell* vs. *Circle-Minus* or *Square-Moon* and *Heart-Bell* vs. *Square-Phone* indicate that if one shape feels thinner than the other, they would be easier to tell apart as suggested by Hypothesis 6. *Heart-Bell* or *Circle-Heart* vs. *Circle-Star* indicate that if such volume difference is not present, extreme feature differences can help. Those patterns align with our participants' reports about features that complicate icon distinguishability. They explicitly mentioned similar bodies and small feature differences 19 times, and similar volume 6 times. However, we cannot fully confirm those patterns, i.e. Hypothesis 6, due to the small number of significant effects we found, and since shapes like *Star-Bell*, *Heart-Phone*, or *Plus-Bell* performed only acceptably. Interestingly, participants also mentioned that different feature numbers and locations helped them distinguish the shapes. However, we would then have expected *Triangle* to be less scattered throughout the *icon distinguishability* ratings as its feature distribution and their number are clearly different compared with most other shapes. In conclusion, we did not find clear patterns for the distinguishability of icons; however, we observed that participants found our icon set mostly acceptable to distinguish, with 93% of our icon pairs receiving ratings higher than 5 and 60% higher than 6 of 7 points.

People do not benefit from feature properties of neighbored icons. One primary purpose of this study was to identify whether people use information from a neighbored icon to recognize features better and improve icon recognition. Overall, we gathered high ratings for *icon distinguishability*, indicating that individual icon recognition has a more dominant influence on the usability of textile user interfaces. From Figure 7, we clearly see that within icon pairs, recognition of individual icons does not vary considerably; for example, *Moon* was recognized similarly well independent of whether it was placed next to a *Square* or *Star*. In contradiction to Hypothesis 3, this shows that our participants gained no benefit from the feature differences between the icons. This surprised us as we informally observed many participants—especially in the beginning—switching back and forth between the icons or even palpating them in parallel using multiple fingers. Therefore, we expected participants to better recognize features such as angled edges, slight curvatures, or merging corners. Also, when comparing ours and results from [31], we see no benefits from recognizing multiple icons together. Our average recognition times of 4.85 to 28.83 seconds (M=20.98*s*, SD=3.35*s*) match their results of (M=7.84*s*, SD=1.68*s*) considering that two icons were palpated, that were more distant from the homing position, and that were unknown to the user. Thus, we found that next to Hypothesis 3, Hypothesis 2 does also not hold.

5 Limitations and Future Work

For our user study, we did not familiarize participants with the icons to avoid confounding the ratings for *icon distinguishability* and to limit biases from a participant's mental image of a shape. This, however, led to a challenging analysis: Since we did not share an icon vocabulary with the participants, their responses do not map directly to terms such as "correct", "satisfying", or "incorrect". Therefore, we followed the procedure described in Section 3.5.2. However, having results that align with previous research makes us confident that our results are valid. To make our classification as transparent as possible, we provide our classification data, including machine-translated responses from our participants, in the supplementary material of this paper.

Since our participants did not know the icons, we cannot ensure whether they always answered the questionnaire for the correct icon—although we did not observe any problems in this regard.

Our study observed the potential of complex shapes to create haptic illusions depending on the context. We could validate this in future experiments by testing icon recognition performance when creating icon associations before the icons are palpated. From this, we expect to be able to create more abstract shapes that can be used manifoldly in many contexts and offer good recognizability—although visual recognition may suffer.

We only tested ten different icons in combination. We decided on our icon set since the icons fit the smart home context and were already investigated in a single-icon user study. The smart home context offered a variety of control elements with different shape characterizations to compare. Although we had to limit our icon set to keep the study in a reasonable time, we assume that our results are not limited to our set or even the smart home domain since our findings are based on the general shape features and independent of the domain knowledge. Nonetheless, the number of samples per interesting pair categories (i.e., for example, thin vs. thick icons or pairs with only a few feature differences) was small. While our icon set already provided first insights, repeating our experiment with more icons for such pair categories—especially to further investigate Hypothesis 6—would lead to further insights on haptic recognition and distinguishability and hopefully strengthen our findings and clarify when icon pairs contradicted them. Furthermore, since we could not verify Hypothesis 1, testing other complex shapes is even more promising and could lead to further insights into the recognition and distinguishability patterns we discussed. Such shapes could be similar to those described in Figure 4, compositions of shapes like a sun/brightness shape, or open shapes like emojis or swirls. For studies investigating such icons, we, however, assume that with the number of icon features, icon size also needs to increase to avoid feature blending. Alternatively, as creating salient features was one of the major issues we encountered for our 18 mm large icons—and led to the invalidation of Hypothesis 5—, future research could also consider emphasizing small features, for example, by changing the texture of the features only. We expect this to make icons more understandable without the necessity of increasing the size.

To keep the study duration acceptable, we kept factors like size, orientation, texture, height of our icons, and curvature of the base surface constant. However, we envision that users will typically use those icons on different everyday objects. Therefore, the form factor of the object on which the icons are placed requires those factors to change We expect that especially the orientation and the curvature of the baseline to make recognition difficult as users will be unsure about the icon's ground side and as curvature can deform the icons. Furthermore, we kept the distance between the icons constant in our user study. While this could be done to investigate further Hypothesis 2 and 3, we do not expect that reducing distance would improve recognition as we would argue that the perceived distance of the icons is already small and participants already palpated the icons simultaneously without noticeable problems.

Young computer science students mainly participated in our user study. Since haptic capabilities can change due to factors like user age, larger-scale user studies should follow our work to investigate the impact of these factors. Also, using the non-dominant hand could reduce the recognition capabilities of the users. As we assume non-dominant hand usage will occur in real-world applications, future user studies should also include handedness as a factor.

In the presented user study, we combined our icons without restrictions to get as many feature combinations as possible. Although we expect very different icons to be placed next to each other, for example, in an application selection area on a user interface, we did not provide clues like this to our participants. Thus, the combinations felt somewhat arbitrary in our study. In upcoming experiments, we also want to investigate how recognition changes when participants know that neighboring icons semantically belong together, e.g., with combinations like *Plus*/*Minus* or volume up/down symbols, but also by contextualizing icons using the interface purpose, i.e., that the icons are used for application selection. If recognition improves in such situations—which is what we assume—investigating how recognition times evolve with the number of icons will be interesting to estimate exploration times in textile interfaces.

Overall, we gathered first insights into human capabilities of haptically exploring our icons in our study. However, from a practicability perspective, other factors like icon familiarity and information about their position will significantly improve recognition in most real-world scenarios. We assume that knowing the position and meaning of interface elements at home will improve recognition time, success, and distinguishability as already a subset of shape features will be sufficient to discriminate the palpated icon from the existing possibilities. Even in situations where the interface is unknown, for example, for a guest in a hotel, the user's intention to control an object will lead to assumptions about what the icon should look like and thus improve identification. In the future, we plan to test scenarios like this and observe how people palpate complete interfaces.

6 Conclusion

We investigated the eyes-free distinguishability of textile icons and how recognition changes when other icons are nearby. For this, our participants palpated pairs of icons without knowing their shapes, guessed the icons, and rated how easy they were to tell apart. We tested six hypotheses about icon recognition and how recognition changes if users explore two icons together. Surprisingly, we found that, in general, those did not hold for our set of textile icons. For recognition time and success, except for the worst-performers *Bell*, *Phone*, and *Plus*, we found that most of our investigated icons performed similarly. Our icons were easy to tell apart, although the differentiation patterns, like different shape volumes, did not impact our ratings as much as expected. Together with our finding that neighboring icons do not help users when recognizing icons, our results indicate that interface designers can combine such icons freely as long as the icons themselves provide acceptable recognition rates.

Acknowledgments

This work was funded in part by the German B-IT Foundation, and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant no. 425869111, and is part of the DFG Priority Program SPP 2199, Scalable Interaction Paradigms for Pervasive Computing Environments.

References

[1] Roland Aigner, Mira Alida Haberfellner, and Michael Haller. 2022. SpaceR: Knitting Ready-Made, Tactile, and Highly Responsive Spacer-Fabric Force Sensors for Continuous Input. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 68, 15 pages. https://doi.org/10.1145/3526113.3545694

[2] Daniel Bühler, Fabian Hemmert, and Jörn Hurtienne. 2020. Universal and Intuitive? Scientific Guidelines for Icon Design. In *Proceedings of Mensch Und Computer 2020* (Magdeburg, Germany) *(MuC '20)*. Association for Computing Machinery, New York, NY, USA, 91–103. https://doi.org/10.1145/3404983.3405518

[3] S. Cecchetto and R. Lawson. 2015. Simultaneous Sketching Aids the Haptic Identification of Raised Line Drawings. *Perception* 44, 7 (2015), 743–754. https://doi.org/10.1177/0301006615594695 PMID: 26541052.

[4] Ben P. Challis and Alistair D.N. Edwards. 2001. Design principles for tactile interaction. In *Haptic Human-Computer Interaction*, Stephen Brewster and Roderick Murray-Smith (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 17–24.

[5] Scott Gilliland, Nicholas Komor, Thad Starner, and Clint Zeagler. 2010. The Textile Interface Swatchbook: Creating graphical user interface-like widgets with conductive embroidery. In *International Symposium on Wearable Computers (ISWC) 2010*. Institute of Electrical and Electronics Engineers, Seoul, South Korea, 1–8. https://doi.org/10.1109/ISWC.2010.5665876

[6] Maas Goudswaard, Abel Abraham, Bruna Goveia da Rocha, Kristina Andersen, and Rong-Hao Liang. 2020. FabriClick: Interweaving Pushbuttons into Fabrics Using 3D Printing and Digital Embroidery. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS*

'20). Association for Computing Machinery, New York, NY, USA, 379–393. https://doi.org/10.1145/3357236.3395569

[7] Chris Harrison and Scott E. Hudson. 2009. Providing Dynamically Changeable Physical Buttons on a Visual Display. 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, 299–308. https://doi.org/10.1145/1518701.1518749

[8] Paul Holleis, Albrecht Schmidt, Susanna Paasovaara, Arto Puikkonen, and Jonna Häkkilä. 2008. Evaluating Capacitive Touch Input on Clothes. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services* (Amsterdam, The Netherlands) *(MobileHCI '08)*. Association for Computing Machinery, New York, NY, USA, 81–90. https://doi.org/10.1145/1409240.1409250

[9] Pradthana Jarusriboonchai, Emma Napari, Oskar Juhlin, and Jonna Häkkilä. 2019. Exploring Non-Emissive Wearable Display as a Clothing Accessory. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers* (London, United Kingdom) *(UbiComp/ISWC '19 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 89–92. https://doi.org/10.1145/3341162.3343787

[10] Amy A Kalia and Pawan Sinha. 2011. Tactile picture recognition: Errors are in shape acquisition or object matching. *Seeing and Perceiving* 24 (2011), 1–16.

[11] Pouya M Khorsandi, Lee Jones, Vandad Davoodnia, Timothy J Lampen, Aliya Conrad, Ali Etemad, and Sara Nabil. 2023. FabriCar: Enriching the User Experience of In-Car Media Interactions with Ubiquitous Vehicle Interiors Using E-Textile Sensors. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1438–1456. https://doi.org/10.1145/3563657.3595988

[12] Konstantin Klamka, Raimund Dachselt, and Jürgen Steimle. 2020. Rapid Iron-On User Interfaces: Hands-on Fabrication of Interactive Textile Prototypes. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376220

[13] Lisa J. Larsson, Géraldine Morin, Antoine Begault, Raphaëlle Chaine, Jeannine Abiva, Evelyne Hubert, Monica Hurdal, Mao Li, Beatriz Paniagua, Giang Tran, and Marie-Paule Cani. 2015. Identifying Perceptually Salient Features on 2D Shapes. In *Research in Shape Modeling*, Kathryn Leonard and Sibel Tari (Eds.). Springer International Publishing, Cham, 129–153.

[14] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research Methods in Human-Computer Interaction.* Morgan Kaufmann, Cambridge, MA, USA.

[15] Samuel Lebaz, Christophe Jouffrais, and Delphine Picard. 2012. Haptic identification of raised-line drawings: high visuospatial imagers outperform low visuospatial imagers. *Psychological research* 76, 5 (2012), 667–675. https://doi.org/10.1007/s00426-011-0351-6

[16] Susan J. Lederman and Lynette A. Jones. 2011. Tactile and Haptic Illusions. *IEEE Transactions on Haptics* 4, 4 (2011), 273–294. https://doi.org/10.1109/TOH.2011.2

[17] Fabrizio Leo, Caterina Baccelliere, Aleksander Waszkielewicz, Elena Cocchi, and Luca Brayda. 2018. Tactile Symbol Discrimination on a Small Pin-Array Display. In *Proceedings of the 2018 Workshop on Multimedia for Accessible Human Computer Interface* (Seoul, Republic of Korea) *(MAHCI'18)*. Association for Computing Machinery, New York, NY, USA, 9–15. https://doi.org/10.1145/3264856.3264858

[18] Yiyue Luo, Kui Wu, Tomás Palacios, and Wojciech Matusik. 2021. KnitUI: Fabricating Interactive and Sensing Textiles with Machine Knitting. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 668, 12 pages. https://doi.org/10.1145/3411764.3445780

[19] Sara Mlakar, Mira Alida Haberfellner, Hans-Christian Jetter, and Michael Haller. 2021. Exploring Affordances of Surface Gestures on Textile User Interfaces. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1159–1170. https://doi.org/10.1145/3461778.3462139

[20] Sara Mlakar and Michael Haller. 2020. Design Investigation of Embroidered Interactive Elements on Non-Wearable Textile Interfaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3313831.3376692

[21] Annika Muehlbradt, Gregory Whiting, Shaun Kane, and Laura Devendorf. 2022. Knitting Access: Exploring Stateful Textiles with People with Disabilities. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1058–1070. https://doi.org/10.1145/3532106.3533551

[22] Sara Nabil, Lee Jones, and Audrey Girouard. 2021. Soft Speakers: Digital Embroidering of DIY Customizable Fabric Actuators. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Salzburg, Austria) *(TEI '21)*. Association for Computing Machinery, New York, NY, USA, Article 11, 12 pages. https://doi.org/10.1145/3430524.3440630

[23] Annie W. Y. Ng and Alan H. S. Chan. 2014. Tactile symbol matching of different shape patterns: Implications for shape coding of control devices. In *Proceedings of the international multiconference of engineers and computer scientists*, Vol. 2. 5 pages.

[24] Oliver Nowak, René Schäfer, Anke Brocker, Philipp Wacker, and Jan Borchers. 2022. Shaping Textile Sliders: An Evaluation of Form Factors and Tick Marks for Textile Sliders. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 214, 14 pages. https://doi.org/10.1145/3491102.3517473

[25] Patrick Parzer, Adwait Sharma, Anita Vogl, Jürgen Steimle, Alex Olwal, and Michael Haller. 2017. SmartSleeve: Real-Time Sensing of Surface and Deformation Gestures on Flexible, Interactive Textiles, Using a Hybrid Gesture Detection Pipeline. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 565–577. https://doi.org/10.1145/3126594.3126652

[26] Yuecheng Peng, Danchang Yan, Haotian Chen, Yue Yang, Ye Tao, Weitao Song, Lingyun Sun, and Guanyun Wang. 2024. IntelliTex: Fabricating Low-cost and Washable Functional Textiles using A Double-coating Process. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 859, 18 pages. https://doi.org/10.1145/3613904.3642759

[27] Myrthe A Plaisier, Wouter M Bergmann Tiest, and Astrid ML Kappers. 2009. Salient features in 3-D haptic shape perception. *Attention, Perception, & Psychophysics* 71, 2 (2009), 421–430.

[28] Ivan Poupyrev, Nan-Wei Gong, Shiho Fukuhara, Mustafa Emre Karagozler, Carsten Schwesig, and Karen E. Robinson. 2016. Project Jacquard: Interactive Digital Textiles at Scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4216–4227. https://doi.org/10.1145/2858036.2858176

[29] Emmi Pouta and Jussi Ville Mikkonen. 2022. Woven ETextiles in HCI — a Literature Review. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1099–1118. https://doi.org/10.1145/3532106.3533566

[30] Jeff Sauro and James R Lewis. 2016. *Quantifying the user experience: Practical statistics for user research.* Morgan Kaufmann, Cambridge, MA, USA.

[31] René Schäfer, Oliver Nowak, Lovis Bero Suchmann, Sören Schröder, and Jan Borchers. 2023. What's That Shape? Investigating Eyes-Free Recognition of Textile Icons. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 580, 12 pages. https://doi.org/10.1145/3544548.3580920

[32] Maarten W.A. Wijntjes, Thijs van Lienen, Ilse M. Verstijnen, and Astrid M.L. Kappers. 2008. Look what I have felt: Unidentified haptic line drawings are identified after sketching. *Acta Psychologica* 128, 2 (2008), 255–263. https://doi.org/10.1016/j.actpsy.2008.01.006

[33] Tony Wu, Shiho Fukuhara, Nicholas Gillian, Kishore Sundara-Rajan, and Ivan Poupyrev. 2020. ZebraSense: A Double-Sided Textile Touch Sensor for Smart Clothing. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 662–674. https://doi.org/10.1145/3379337.3415886

A Appendix

The following pages contain additional diagrams providing more details or alternative visualizations of our study results.

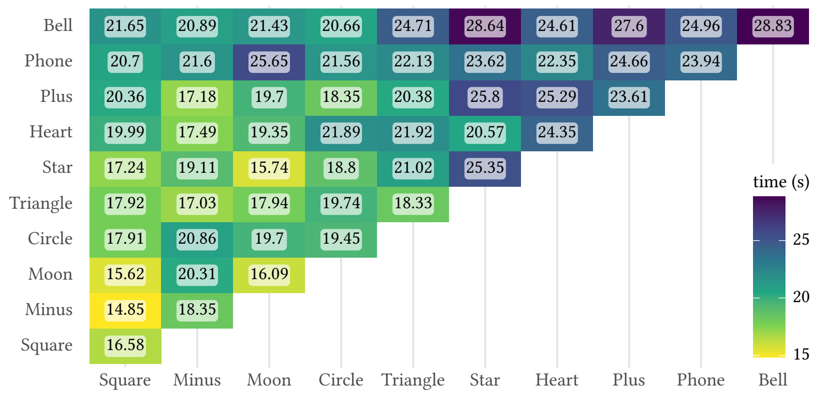

Figure 10: Heatmap showing the average *recognition time* per icon pair. Cells are tinted according to their *recognition time* from yellow (fast) to blue (slow). Especially icon pairs including *Bell* and *Phone* needed a long recognition time, with *Bell-Bell*, *Star-Bell*, and *Plus-Bell* being the worst pairs.

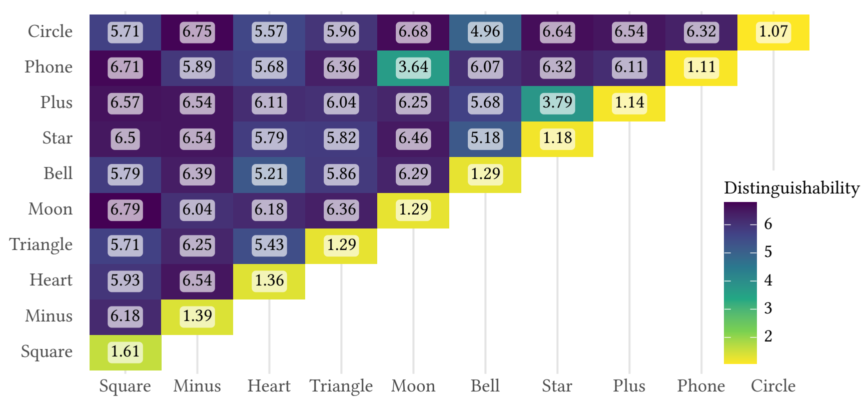

Figure 11: Heatmap showing the perceived haptic distinguishability of the icon pairings, based on the participants' feedback. The higher the ratings, the easier the shapes were to tell apart. Twin icons were rated the lowest and non-twin icons usually received high scores, except for *Moon-Phone* and *Plus-Star*, which, on average, received noticeably lower scores.

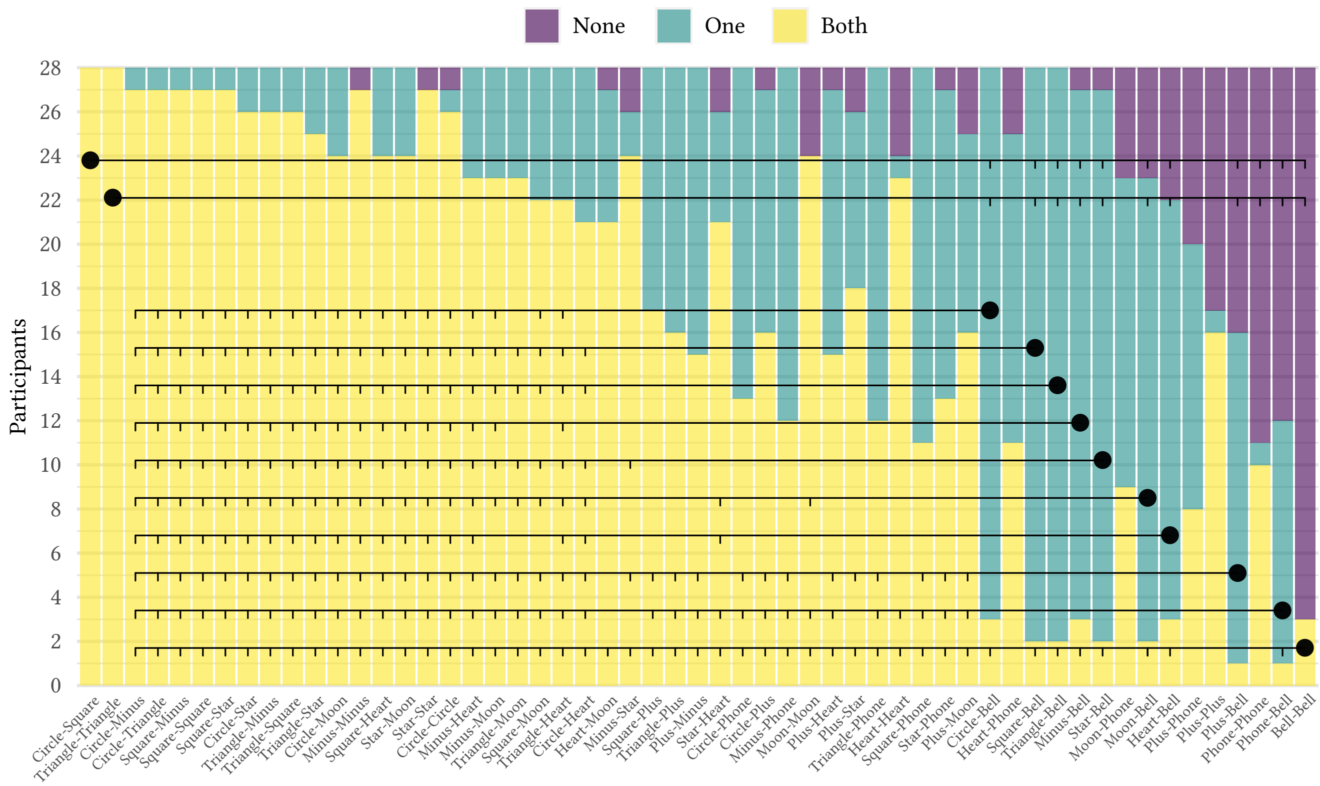

Figure 12: The number of icon pairs that were named correctly by our participants. We differentiate between icon pairs where both (yellow), only one (teal), or no (violet) icon was recognized. Horizontal bars denote significant differences where all icon pairs connected with a tick are significantly different to the pair connected with a circle. All icon pairs containing *Bell* performed significantly worse compared to our best performers, which mostly included simple geometrical shapes.