Tool Extension in Human–Computer Interaction

Joanna Bergström, Aske Mottelson, Andreea Muresan, Kasper Hornbæk
University of Copenhagen

ABSTRACT
Tool use extends people’s representations of the immediately actionable space around them. Physical tools thereby become integrated in people’s body schemas. We introduce a measure for tool extension in HCI by using a visual-tactile interference paradigm. In this paradigm, an index of tool extension is given by response time differences between crossmodally congruent and incongruent stimuli; tactile on the hand and visual on the tool. We use this measure to examine if and how findings on tool extension apply to interaction with computer-based tools. Our first experiment shows that touchpad and mouse both provide tool extension over a baseline condition without a tool. A second experiment shows a higher degree of tool extension for a realistic avatar hand compared to an abstract pointer for interaction in virtual reality. In sum, our measure can detect tool extension with computer-based tools and differentiate interfaces by their degree of extension.

CCS CONCEPTS
• Human-centered computing → HCI theory, concepts and models; HCI design and evaluation methods; Laboratory experiments; User studies; Empirical studies in HCI;

KEYWORDS
Tool extension, body schema, peripersonal space, input

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM Reference Format:

1 INTRODUCTION
Computer systems may be viewed as tools to achieve specific goals and human-computer interaction (HCI) thereby seen as tool use [e.g., 23, 30]. Although computer-based tools are complex, under this view findings about tool use may be applied to HCI. We investigate if the advances during the past 20 years in understanding the influence of tools on our perception of our bodies, called tool extension, may apply to HCI, too.

Tool extension was first empirically characterised by Iriki and colleagues [24]. They trained macaque monkeys to retrieve objects using a rake. When so trained, the neural areas associated with the hand also registered visual activity around the rake and the area reachable with the rake. Since Iriki’s study, these findings have been extended to humans, showing a similar influence of manual tools on our perception and thinking. In particular, the use of a tool has been shown to change our motor control [11] and perception around the tool [3]. Recent reviews provide further details on how tools act as extensions of our bodies, modify the representations of our peripersonal space, and incorporate into our body schema [6, 20, 27].

The methodologies used to establish tool extension, and their resulting findings, have only rarely been applied to computer-based tools. In one exception, Bassolino and colleagues [3] used the audio-tactile integration paradigm to find that use of mouse extends peripersonal space representation. With the audio-tactile paradigm, extension shows as a lack of difference between response times to tactile stimuli during concurrent sounds near or far to the hand holding the tool. Therefore, it cannot be used to assess degrees of tool extension.

In another exception, Sengül and colleagues [35] employed virtual tools, but merely replicate tool extension findings instead of attempting to use tool extension to compare ways of interacting. Despite these exceptions, researchers in neuropsychology have lamented a lack of studies on more complex tools [16]. We are similarly unaware of previous studies having applied the methodologies for studying tool extension in HCI. The associated findings have not been used to evaluate specific interaction styles, although mentioned in discussions of embodied interaction [e.g., 25].

We introduce a measure of tool extension to HCI, and validate it in two studies. The first study demonstrates that the visual-tactile interference paradigm shows tool extension with two common input devices: a mouse and a touchpad. By comparing the mouse and the touchpad to a condition where no input is given, we show that the paradigm applies beyond manual tools to computer-based tools. The second study shows that minute details of the interaction styles may change the degree of tool extension. We find that the presence or absence of an avatar’s arm provide different degrees of extension.

Thereby, we introduce a measure of tool extension to HCI that can quantify claims about embodiment of computer-based tools. We contribute a validation of this measure by showing that it can be used to characterize different interaction styles.
2 RELATED WORK
We first review work on tool extension from neuroscience and psychology, which have established the influence of using tools on our perception of the space around us. Those studies have mainly been done using physical tools. Nevertheless, a few studies have investigated computer-based tools; we review those as well. Finally, we discuss how the methodology used in these studies can be made usable in HCI.

Tool Extension in Neuroscience and Psychology
Both neuroscience and psychology have been concerned with understanding tool use in general and in particular how tool use affects our perception of our body and the environment around us (for reviews, see [6, 20, 27]). Whereas the idea that our representation of our body (often called body schema) is affected by tool use is more than a hundred years old [22], the study by Iriki and colleagues was a key, modern introduction of this idea [24]. They showed that the brains of monkeys, in the intraparietal cortex which integrates visual and somatosensory information, are affected by tool use in a surprising way. Before using a rake (and when passively holding it), bimodal neurons in monkeys’ brains only responded to visual stimuli near the hand. However, after using the rake, brain activity was also seen for stimuli along the tool. Thus, the tool seems to be incorporated into the body.

It was later shown that tools affect humans similarly, using a behavioural task, rather than neural measures [28]. Maravita and colleagues used toy golf clubs with vibrotactile stimuli at their handles (in an upper and a lower location) and visual distractors at their tips (also in an upper and lower location). Participants had to judge the location of the vibrotactile stimuli while ignoring the visual stimuli. This setup is based on the visual-tactile interference paradigm, also known as the crossmodal congruency task [37]. This task assumes that stimuli in different modalities (e.g., tactile and visual) interact; in particular, stimuli that are incongruent, say visual in upper location and tactile in lower, lead to higher response times (or lower accuracy) because they require effort to ignore. In the study by Maravita, participants held two golf clubs and indeed showed a crossmodal congruency effect. Crucially, this effect changed when participants cross the golf clubs so that it depended on the tip of the held club rather than on the tip closest to the hand. Despite the complexity of the setup, it shows that tools appear to integrate with our hands.

This finding has been extended and validated in several ways. Besides the crossmodal congruency task, the effect has been shown for length estimations. For instance, tool use makes people perceive their forearm as being longer than before tool use [11], and merely observing tool use compresses the perception of distance [7]. Targets are also perceived as nearer in space when holding a stick able to reach them, compared to when participants do not hold a stick [39]. Tool extension has also been shown for mechanical grabbers [11], rakes [17], canes [36], and wheelchairs [15].

These findings are bounded in important ways. Practice with the tool is critical. Several studies show that if you are highly practiced with a tool, merely holding it is enough to cause the tool extension to occur. In a study of tennis rackets, for instance, holding your own racket is enough for tool extension whereas generic rackets require use for tool extension [5]. The tip of the tool also seems particularly relevant, whereas non-active parts of the tool do not seem to be incorporated into the body schema [21].

Tool Extension from Computer-based Tools
In a review of tool extension, Holmes and Spence [22] distinguished physical tools (e.g., rakes, golf clubs; used in the studies just cited) from pointing tools (e.g., laser pointers, the mouse), and detached devices (e.g., telesurgical devices). The latter two categories are of relevance to HCI because of their resemblance to input techniques in HCI; however, both in the review as well as in recent work, these two categories are rarely explored.

In the second category, Bassolino and colleagues [3] presented a study of whether using a computer mouse provides tool extension. Bassolino used an audio-tactile interaction paradigm in which sounds appeared near the hand and near the pointer controlled by the mouse (70 cm away from the hand). The assumption in this setup is that audio from far and from near might interfere with response times to a tactile stimulus on the hand in different ways. In particular, far sounds should increase response times only if the far space is considered separate from near space (thereby this paradigm is different from the crossmodal congruency task). The study results show that when not using a mouse, the response time to tactile stimuli with sounds near the hand is lower compared to far sounds, whereas when either holding or actively using a mouse, there is no difference between those response times. Bassolino and colleagues interpret this as showing extension of peripersonal space by the mouse. However, their study concerns the familiarity of the used tool (the mouse) and does not intertwine real-world human-computer interaction with a measure of tool extension.

In the third category, a few studies have recently experimented with applying tool extension paradigms to other types of computer-based tools [8, 35]. Sengül and colleagues [35] transferred earlier findings on tool extension with golf clubs (reported in [28]) to virtual reality (VR) in which participants used a haptic device. However, their findings are mainly about replication of the original study in VR and not about interaction styles therein.

Physical tools (Holmes and Spence’s first category) are relevant in HCI, too, as they might be coupled with computer-based tools. We are unaware of work on tangible or physical interfaces using tool extension methodology. Yet, the study by Alzayat and colleagues [1] is relevant here. They studied an object manipulation task using a physical and a virtual tool. While doing the manipulation task, participants had to
Issues in Transferring Tool Extension to HCI

Tool extension is often discussed as a concept that is important to HCI [e.g., 18, 25]. However, as seen above, few studies use this concept for computer-based tools; none within HCI. We believe this is due to three issues in adapting the methodology to provide a useful measure for HCI.

One issue in transferring tool extension methodology is that HCI studies require participants to actually use the tool or user interface, beyond merely completing the tasks necessary for measuring extension. Tool use rarely happens in existing work. For instance, in the paper by Bassolino and colleagues [3], the description of the experiment leaves it unclear what participants do with the mouse, in particular, if they do any real-world interactions, such as aiming and selecting. Similarly, neuropsychology has used tool extension to show inclusion of artificial limbs within the body schema (e.g., in the Rubber Hand Illusion [32]). The Rubber Hand Illusion, however, does not require action but merely that participants experience visuo-tactile synchronization to induce tool extension. With physical tools in neuroscience, action has been incorporated by small movements, such as crossing the tools, instead of using them for their real functions (e.g., clubs for golf [28]). For HCI, however, we need to find a way to intertwine the task for measuring extension and the tool-specific interactions. Doing so is critical because action is widely believed to be essential for the representation of peripersonal space [26] and, of course, a key in human-computer interaction.

A second issue is to apply the findings from physical tools to computer-based tools. In particular, tool extension has been studied mainly for physical tools that are grasped. The role of active and passive grasp is widely studied, as is the effect of previous experience in using a particular tool. From the literature, however, it remains unclear if tools that are not grasped, such as a computer touchpad, can induce tool extension. Further, the findings from psychology and neuroscience on the differences between an empty hand and a tool are confounded in that we found no examples where the empty hand functions as a tool. For HCI, however, many interactions do not require grasping, and in many cases an empty hand works as the tool (as it would, for instance, in mid-air pointing or non-controller interaction in VR).

A third issue is to transfer the stimuli for the crossmodal congruency task from physical tools to computer-based tools. Experiments on visual-tactile congruency are often done with very particular hand orientations and modes of grip (e.g., holding a golf club with a thumb up). With physical tools, the experiments use congruency between the stimuli within a single plane and thus obtain a direct spatial mapping (e.g., thumb up for tactile stimulus and the same upper side of the grasped golf club for a visual stimulus). Computer-based tools, however, pose a challenge to that mapping. For instance, movements of a mouse on a table plane map to cursor movements on a display, which is often positioned on a plane orthogonal to the table. But how to map the tactile stimuli on the index finger and thumb to the cursor on the display? Earlier work has not faced this problem. Therefore, it remains unclear if we can provide visual and tactile stimuli on those two distinct planes in a way that make tool extension methodology applicable to computer-based tools and HCI.

We find that these three issues are the reason why tool extension methodologies have not been used in HCI. Transferring methods from neuroscience and psychology often requires adaptation work when applied in HCI. The sense of agency, for instance, plays a vital role in neuropsychology, and was initially transferred to HCI by Coyle and colleagues in 2012 [13]. After that, several papers have been required to further test it [4] and extend it to work in actual task performance, for instance in VR [12]. In the experiments that follow, we first show that the three issues discussed can be dealt with and that this methodology allows HCI to reason about computer-based tools in new ways. But first, we introduce the basic idea of a tool extension measure in HCI.

3 A MEASURE OF TOOL EXTENSION FOR HCI

We introduce a measure of tool extension to HCI by applying a visual-tactile interference paradigm. The purpose of our measure is to clear the stumbling blocks discussed in the previous section.

The visual-tactile interference paradigm is suitable for HCI, because most user interfaces rely on visual output and feedback. Therefore, this paradigm allows intertwining our measure with interaction, such as with controlling a cursor on a display. Figure 1 presents the procedure for the visual-tactile interference paradigm, using the interfaces and the setup of our Experiment 1 as examples.

In the visual-tactile interference paradigm, participants make speeded discriminations between two tactile stimuli (e.g., up/down) in the presence of a congruent (e.g., up/down) or incongruent (e.g., down/up) visual distractor stimulus. The visual distractors can be embedded on the visual feedback of the user interface, such as above or below a cursor (Figure 1). The actuators for tactile stimuli are placed in a way which can be mapped to the visual stimuli. For example, in interaction with a mouse, forward movements of the hand are mapped to upward movements of a cursor. As the posture of the hand controlling a mouse or a touchpad entails the index finger tip being located more forward than the thumb, the actuators
Discriminate between the vibrotactile stimuli (in the presence of a visual distractor)

Move the cursor to select a target circle (or observe it move)

Move the cursor to the next target circle (or observe it move)

Figure 1: The procedure in our measure using the visual-tactile interference paradigm. The procedure uses the setup of our Experiment 1 to exemplify the role of input interfaces and the visual feedback. However, the procedure applies beyond this particular setup with modifications according to the spatial configurations of the interface.

Discriminate between the vibrotactile stimuli (in the presence of a visual distractor)

Response time

Move the cursor to select a target circle (or observe it move)

Move the cursor to the next target circle (or observe it move)

Figure 1: The procedure in our measure using the visual-tactile interference paradigm. The procedure uses the setup of our Experiment 1 to exemplify the role of input interfaces and the visual feedback. However, the procedure applies beyond this particular setup with modifications according to the spatial configurations of the interface.

Responses to tactile stimuli are given with a modality that has minimal interference with conducting the experiment task or interacting with the interface. For example, in our experiments the participants gave input with a hand, and therefore we used foot pedals to record responses (which are also common in previous studies on visual-tactile interference [e.g., 28, 35]). Foot pedals can be used, for instance, with one foot to respond to a tactile stimulus on the index finger by lifting the toes and to a stimulus on the thumb by lifting the heel (Figure 1). Another way to record responses to tactile stimuli is to use key presses with the hand not stimulated, for instance by responding with left and right keys to tactile stimuli on the index and middle finger [e.g., 40]. To tightly intertwine interaction with the measure, a discrimination trial of tactile stimuli is performed after each instance of action (e.g., selecting a target with the cursor, as Figure 1 and in our Experiment 1).

Our measure for tool extension in HCI is crossmodal congruency effect (CCE). CCE is validated in studies ranging from monkey’s visuo-somatosensory neural activity [24] to behavioural measures with brain-damaged patients [27]. It has been shown to increase for objects that can be easily integrated in body representation (e.g., in the Rubber Hand Illusion [32, 40]), and it reflects changes in peripersonal space representation after tool use [e.g., 20, 27, 28]. CCE is defined as the difference between mean response times (RTs) in incongruent conditions and congruent conditions within the visual-tactile interference paradigm. Figure 2 presents the four stimuli conditions, and the correct responses to tactile stimuli, again using the setup of our Experiment 1 as an example. CCE is calculated from RTs measured in repetitions of each of the four congruence conditions as presented in Figure 2.

With tool extension, the RTs in incongruent conditions are longer than in congruent conditions. Therefore, tools that are easily integrated to the peripersonal space or body representation result in increased CCEs compared to tools that are not. Thereby, our measure allows comparison of interaction styles by the degree of tool extension they provide.

4 EXPERIMENT 1

The purpose of the first experiment is to investigate if our measure shows tool extension with computer-based tools. For this purpose, we use two common input devices: a mouse and a touchpad. We also include a passive condition as a baseline, where no input device is used.

A mouse has previously suggested to provide tool extension with an audio-tactile interaction paradigm. Finding extension by using a mouse in this experiment can therefore validate the suitability of our measure for tool extension.

<table>
<thead>
<tr>
<th>Vibrotactile stimulus</th>
<th>Visual distractor</th>
<th>Correct responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruent conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On the index finger</td>
<td>Above the cursor</td>
<td>Lift the toes</td>
</tr>
<tr>
<td>On the thumb</td>
<td>Below the cursor</td>
<td>Lift the heel</td>
</tr>
<tr>
<td>Incongruent conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On the index finger</td>
<td>Below the cursor</td>
<td>Lift the toes</td>
</tr>
<tr>
<td>On the thumb</td>
<td>Above the cursor</td>
<td>Lift the heel</td>
</tr>
</tbody>
</table>

CCE = Mean RT in Incongruent Conditions - Mean RT in Congruent Conditions

Figure 2: The stimuli and responses for incongruent and congruent tactile and visual stimuli, and the calculation of the CCEs based on response times (RTs).
Tool extension has previously been shown mainly with manual tools that are grasped. We use a touchpad to examine whether findings on tool extension apply to tools that are not grasped, as such input interfaces are common in HCI.

In this experiment, we also intersperse interaction and our measure for tool extension. For example, the participants use the mouse and the touchpad to move and select with a cursor in a setting wherein we do not control their grip and hand posture with the input devices. We restrain from using other means for high control seen in previous studies, such as chin rests. By relaxing some of the control in the experiment, we examine if our measure can show tool extension when integrated with human–computer interaction.

Participants
Data were collected from 12 right-handed participants (7 identified themselves as females and 5 as males, with an mean age of 25.92 ± SD of 3.85 years). Six of them used a mouse and ten of them a touchpad daily.

Task
The experiment task was to discriminate between two locations for vibrotactile stimulus while ignoring visual distractor stimuli. The tactile stimuli were presented either on the index finger or the thumb. The response to a tactile stimulus was given with a foot pedal by lifting the toes to indicate a stimulus on the index finger or lifting the heel to indicate a stimulus on the thumb.

Design
The experiment followed a two-by-three within-subjects design with the congruency between the tactile stimuli and the visual distractor as the first factor and input condition as the second.

In two active input conditions, the participants used a mouse or a touchpad to control a cursor on a computer display (Figure 3). In the passive condition, no input device was held and no input was given – the cursor was only observed to move.

The visual distractor stimuli were presented by highlighting either of two dots above or below that cursor. The congruent and incongruent conditions consisted of visual stimuli above or below the cursor with a tactile stimuli on the index finger or thumb, as presented in Figure 2.

The order of the input conditions was counterbalanced. Each of the four congruency conditions were presented 20 times with each input condition in a randomised order. The experiment thus consisted of 360 trials in total (3 input devices × 4 congruency conditions × 20 repetitions). The experiment lasted approximately 30 minutes. The participants received small gifts as a compensation for their time.

Procedure
The participants were instructed to fixate at the cursor at all times, and to respond to the tactile stimulus as quickly as possible by lifting either of the foot pedals, while ignoring the visual stimulus.

Before starting the experiment and before starting each input condition, the participants were asked to lean back on a chair, and to rest their foot on the pedals and lift the heel and toes a few times to ensure a firm enough press to not activate pedals accidentally, and a relaxed and comfortable posture for responding to stimuli. The participants were then reminded of their task.

In the active conditions with input devices, the participants were also instructed to move a cursor around the window and to click with a mouse or tap with a touchpad to get used to their movement-display ratio and the feel of a selection. They were instructed to tap, not click, the touchpad. The procedure is presented in Figure 1. In the active conditions, the participants started the experiment by moving the crosshair cursor to a target circle on the display. There were two target circles, on the left and on the right, and the previous target circle acted as the starting point for the next trial. In the passive condition, the participants did not control the cursor, but only...
observed it move to the targets. The movement speed of the cursor was the average speed recorded in pilot studies.

After successfully selecting the target or observing the cursor to reach the target (in the passive condition), the target disappeared. Then, both a tactile stimulus and a visual distractor stimulus were presented. The visual distractor stimulus on either of the dots above or below the cursor was presented 900 ms after the click on a target circle. With this delay, we aimed to balance between having a short enough delay after moving the cursor to maintain active tool use (in active conditions), but long enough to avoid possible effects of sensory attenuation on perception of the tactile stimuli. Such attenuation may occur when the body part receiving a stimulus is moved during or immediately before the stimulus. Similar delays have been used in previous work [e.g., 40].

The tactile stimulus was given 100 ms after the visual distractor stimulus, because previous work has shown that such a delay maximises the CCE. The duration of both the visual and tactile stimuli were 200 ms (therefore being simultaneously elevated for 100 ms). These durations, too, are similar to previous work, which vary the duration from 100 ms [e.g., 35] to 250 ms [e.g., 28, 40].

The participants responded to the tactile stimulus by lifting a foot pedal. The participants were instructed to not move the cursor (in the active conditions) before responding to stimulus and pressing a foot pedal back down. After a foot pedal was pressed back down, the next target circle appeared on the screen. The participants were allowed to take a short break and rest their foot between the input conditions.

Setup
The experiment setup is presented in Figure 3. The experiment interface was presented on a full screen window on the display. We choose to use a crosshair as a cursor due to its neutral appearance, so as to avoid presenting unrelated directional cues that, for instance, an arrow cursor could have given. The dots for visual distractor stimuli were 100 pixels in diameter and 60 pixels above and below the cursor. The target circles were 60 pixels in diameter at a vertical midpoint and at a horizontal distance of a 1/6 of the width of the window from the left and right sides.

The participants sat in front of a desk. The display was placed on the desk one meter away from the back of the chair participants sat on. The participants were asked to lean back on the chair to keep the distance from their body to the display constant. The input devices and the hand in the passive condition were placed and used within a marked 30 cm wide and 20 cm deep area. The nearest edge of that area was set to a 25 cm distance from the display. The foot pedals were Andoer Game Foot Switch USB connected pedals, attached to the floor underneath the participant’s heel and toes. The foot pedals were located on the floor in relation to the chair in a way that the participant’s right knee would form a maximum of 90 degree angle when the right foot would rest on the pedals.

This posture was chosen to maximise the press on the pedals with a relaxed foot, in order to avoid accidental lifts of pedals.

Apparatus
For tactile stimuli, we used shaftless 3V, 60mA, 13000±3000 rpm coin vibration motors with a diameter of 10 mm and a thickness of 3.4 mm. The vibration motors were controlled with an Arduino Uno. The vibration motors were placed on the medial phalanx of the index finger and proximal phalanx of the thumb. The motors were attached firmly on the fingers with a slightly flexible medical self-adhesive bandage.

The experimental interface for visual stimuli, for controlling the vibration motors, and for logging the foot pedal responses was developed with Processing. The interface was presented on a Samsung SyncMaster 244T LCD display with a 1920 × 1200 resolution. The experimental software ran on a MacBook Pro (2013). The input devices were an Apple Magic Mouse and an Apple Magic Trackpad 2. The tracking speed of both the mouse and the touchpad was set to the third fastest level on the MacBook’s system preferences.

Results
Figure 4 and Table 1 show the results from Experiment 1. For analysing the RTs, trials with an incorrect response to the tactile stimuli were discarded (the error rates are presented in Table 1). Trials with RTs smaller than 200 ms or larger than

<table>
<thead>
<tr>
<th>Mouse</th>
<th>Touchpad</th>
<th>Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>Errors</td>
<td>RT</td>
</tr>
<tr>
<td>Congruent</td>
<td>674.14</td>
<td>1.46</td>
</tr>
<tr>
<td>Incongruent</td>
<td>829.15</td>
<td>7.08</td>
</tr>
<tr>
<td>CCE</td>
<td>155.01</td>
<td>5.63</td>
</tr>
</tbody>
</table>

Figure 4: The average CCEs for the input conditions in Experiment 1. The error bars denote .95 confidence intervals.
1500 ms were also discarded (3.02% of trials with the mouse, 2.19% with the touchpad, and 1.88% in the passive condition) as those represent either anticipations or omissions of stimuli (these procedures in analysis follow previous studies, [e.g., 20, 28, 35]).

We used Bonferroni corrected paired sample t-test to analyse differences between each input condition and the passive condition by using the mean values per condition for each participant. The t-tests show a significant difference between the two input condition pairs: mouse and passive ($t(11)=3.66$, $p=.004$), and touchpad and passive ($t(11)=4.72$, $p<.001$).

Finding a higher CCE with a mouse compared to the passive condition confirms that our measure can show tool extension. Finding a higher CCE with a touchpad compared to the passive condition shows that also tools that are not grasped can extend representation of peripersonal space. Together, these findings also show that our measure for tool extension can be intertwined with human-computer interaction in user experiments.

5 EXPERIMENT 2

The primary purpose of the second experiment is to examine whether our measure can distinguish degrees of tool extension for different interaction styles. Tool extension measures changes in the representation of the actionable space (i.e., body schema), independently of whether the tool is physical or a virtual projection of the body (as shown with neural measures with monkeys [27]). Therefore, another purpose is to examine whether our measure also shows tool extension in free-hand interaction (not holding nor touching an input device). Hands are important tools in HCI, as they are often used for input, for instance, with mid-air, gesture, augmented reality, and VR interfaces. Here, too, we intertwine interaction and our measure for tool extension; this time in a VR setting.

We chose to vary the appearance of an avatar hand in three conditions presented in Figure 5: A realistic hand, a disconnected hand, and an abstract appearance. These appearances were chosen because previous findings on tool extension suggest that appearance of artificial bodies can influence the magnitude of the CCE.

First, we chose to use a realistic avatar hand with an arm as one appearance condition. We hypothesize that a realistic hand shows a highest degree of tool extension, because previous work suggests increased CCE and inclusion of artificial limbs (e.g., in the Rubber Hand Illusion [32, 40]) in body representation.

As a second appearance condition we used a disconnected hand. CCE has been presented as an objective predictor of ownership [40]. Perez-Marcos et al. [33] showed that subjective ownership depends on connectivity of a virtual arm, but did not find such dependency on the alignment of a virtual hand that was connected. Therefore, a disconnected hand was hypothesized to show lower CCE than a realistic hand connected to an arm. The realistic hand with an arm was gender matched for participants, as gender has been shown to influence body ownership of avatars [34]. In contrast, the disconnected avatar hand had a static, neutral appearance.

As a third appearance condition we used an abstract cursor. This appearance was included because previous work has found differences between natural appearances and abstract appearances. For example, Pavani et al. [31] found a larger CCE with stimuli next to a natural shadow of a real hand, compared to an abstract, polygonal shaped shadow. In another study, Pavani et al. [32] showed with rubber gloves that having the gloves visible increased CCE compared to not having gloves at all (only distractor lights). We derive the abstract appearance from these studies: using only spheres at the locations of the fingertips.

Participants

Data were collected from 18 right-handed participants (9 identified themselves as females and 9 as males, with an mean age of 24.39 ± SD of 3.24 years). Twelve of them had tried a VR headset before.

Task

The experiment task was the same as in Experiment 1.

Design

The experiment followed a two-by-three within-subjects design with the congruency between the tactile stimuli and the visual distractor as the first factor and appearance condition as the second. The three appearances were a realistic
Hand and arm, a realistic disconnected hand (without an arm), and an abstract condition wherein only white spheres at the locations of fingertips were visible (Figure 5). The spheres followed participants’ fingertip movements, thus inducing body ownership by visuo-motor synchronisation, which has been shown to correlate with tool extension [31]. All conditions in this experiment were active.

The visual distractor stimuli were presented by highlighting either of white spheres. The congruent conditions consisted of visual stimulus on the white sphere at the fingertip corresponding to the finger receiving the tactile stimulus. The incongruent conditions consisted of the opposite.

The order of the input conditions was counterbalanced. Each of the four congruency conditions were presented 20 times with each input condition in a randomised order. The experiment thus consisted of 360 trials in total (3 input devices × 4 congruency conditions × 20 repetitions). The experiment lasted approximately 30 minutes. The participants received small gifts as a compensation for their time.

Procedure

The participants were instructed to look at the spheres at all times, not to close their eyes, and to respond to the tactile stimulus while ignoring the visual stimulus as quickly as possible by lifting the corresponding foot pedal.

Before starting the experiment and before starting each appearance condition, participants were asked to rest their foot on the pedals and lift the heel and toes a few times to ensure a firm enough press to not activate pedal accidentally, but also a relaxed and comfortable posture for responding to stimuli. The participants were then reminded of their task.

The participants rested their elbow on a table, kept their hand in the air and pointing forward, and their index finger and thumb opened apart so as to keep the spheres on top of each other (Figure 6). Other fingers were kept relaxed (slightly flexed).

All appearance conditions consisted of the same interaction: Moving the hand between two spotlights on the right and on the left. After holding the spheres under a spotlight and not moving the fingers (centroid between them) more than 0.5 cm for 2 seconds, the visual distractor stimulus was presented, followed by a tactile stimulus a 100 ms later. The durations of both visual and tactile stimuli were 200 ms, the same as in Experiment 1.

The participants responded to the tactile stimulus by lifting a foot pedal. The participants were instructed to not move their hand before responding to the stimulus and pressing a foot pedal back down. After a foot pedal was pressed back down, the next spotlight appeared in their view. The participants were allowed to take a short break and rest their foot between conditions.

Apparatus

The tactile stimuli and foot pedal setup were identical to those used in Experiment 1. In addition, we used an HTC Vive with Optitrack motion capture system for finger tracking and VR. The VR scene was created in Unity 2018 and ran on a laptop PC (2.8GHz Intel i7, 16GB RAM, NVIDIA GTX 1060), running Windows 10 Pro. Finger tracking was done with eight Prime13 cameras at a 120 Hz frequency, placed around a table where the study was conducted. We tracked the thumb and index finger using reflective markers attached to a 3D printed frame (Figure 6). We used FinalIK as the inverse kinematic model for more realistic hand and arm movements.

Results

Figure 7 and Table 2 show the results from Experiment 2. As in Experiment 1, trials with an incorrect response to the tactile stimuli were discarded from the analysis of the RTs (those error rates are presented in Table 2). Trials with RTs smaller than

<table>
<thead>
<tr>
<th>Condition</th>
<th>Realistic</th>
<th>Disconnected</th>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruent</td>
<td>710.45</td>
<td>1.25</td>
<td>713.00</td>
</tr>
<tr>
<td>Incongruent</td>
<td>891.74</td>
<td>5.42</td>
<td>872.75</td>
</tr>
<tr>
<td>CCE</td>
<td>181.29</td>
<td>4.17</td>
<td>159.75</td>
</tr>
</tbody>
</table>
200 ms (anticipations) or larger than 1500 ms (omissions) were also discarded (5.21% of trials with the realistic arm, 4.79% with the disconnected hand, and 2.50% in the abstract condition).

We used Bonferroni corrected paired sample t-tests to analyse differences between each appearance condition by using the mean values per condition for each participant. The t-tests show a significant difference between one appearance condition pair: a realistic arm and an abstract condition ($t(17)=3.10, p = .009$). The error rates did not show significant differences.

Finding a higher CCE with a realistic avatar hand compared to the abstract appearance confirms that our measure can distinguish degrees of tool extension. This finding together with the average CCEs is in alignment with previous studies on tool extension: characteristics of avatar hands similar to those applied to shadows [31] and rubber hands [32, 40] influence tool extension in similar ways. For example, natural appearance increases CCE over abstract appearance, and a disconnected hand decreases CCE compared to a hand connected to the body with an arm. Our findings also show that tool extension can be found in free-hand interaction with computer-based tools. Together, these findings imply that our measure shows tool extension for avatars, and can be intertwined with interaction in VR.

6 DISCUSSION
Extending our action space by a tool is central to pointing and manipulating the world. Tool extension is an established concept in psychology and neuroscience about how a person’s body schema may change as a result of tool use. Previous work has claimed it relevant to human-computer interaction (HCI), but so far the empirical methods for establishing tool extension has not been applicable to HCI. Our experiments show that tool extension can work in HCI: We have presented a measure of it that may be collected during interaction, without restrictions in grip or movement. Moreover, the measure helps discriminate levels of extension, for instance between mouse, touchpad, and a passive condition, as well as between a full representation of an avatar arm in VR and an abstract pointer. Thereby, we have created a measure for use in HCI that quantifies a central aspect of computer-based tools, namely whether they incorporate into our body schema. Next, we discuss details of the studies and their implications.

Our Measure for Tool Extension
The key contribution of this paper is introducing and validating a measure of tool extension for HCI. This measure is objective but captures qualities of interaction that HCI has talked about in subjective terms for decades. Those terms include ‘acting through the interface’ [9], ‘ready-to-hand’ [1], ‘transparent’ [2], ‘embodied’ [25], and ‘natural’ [38]. Tool extension does not help quantify the entirety of these qualities, but it does attempt to capture parts of those. As such, it represents a milestone in working quantitatively with the experience of user interfaces.

The adaptations in the present paper are substantial over methodology in psychology and neuroscience. First, our measure does not require chin rests, fixed arms, or unchanging grips. Relaxing these constraints allows using the measure, for instance, in VR and with devices that require a grip (mouse) or not (touchpad). Second, users may do interactive tasks (rather than just the discrimination task as in most studies on tool extension [e.g., 28]). Thereby, the measure can be intertwined with task performance.

The validity of the methodology has been shown mainly as discriminant validity. It was able to distinguish active (mouse, touchpad) and passive conditions in Experiment 1 (as expected from research on physical tools [5]). It was also able to distinguish degrees of tool extension with variants of VR avatars in Experiment 2 which subjective body ownership reports from earlier work have suggested dissimilar [19, 29]. The effect size there was medium (Cohen’s $d=0.548$) for the difference between realistic and abstract appearances.

We find that visual–tactile interference has some benefits for HCI over audio–tactile interaction paradigm, as used by Bassolino and colleagues [3]. First, the visual modality is dominantly used for output and feedback in user interfaces. Most input is mapped to visual output, from keyboard events to text and mouse or hand movements to pointers. Therefore, visual distractor stimuli are easy to embed in real-world interactions. Second, the audio–tactile paradigm requires physical distance from the tool to the interface. Therefore, it may not apply to use with interfaces for VR. Third, the audio–tactile interaction paradigm shows extension of peripersonal space with a tool when no difference is found between response times between the near and far sounds concurrent with the tactile stimuli. As the measure of extension is based on finding no difference, it requires a passive condition as a baseline wherein a difference is found. Fourth, as the audio–tactile paradigm is indeed based on indifference in response times, it cannot be used
for comparing interaction styles. In contrast, our measure is based on the magnitude of the CCE, and can thus be used for comparing the degree of tool extension between interaction styles, as we did in Experiment 2.

Limitations and Open Questions
Our study has several limitations. Most importantly we did not quantify prior experience with the tools tested. Earlier work has shown strong effects due to familiarity [e.g., 5] and we suspect this could moderate our findings, too. Future work should attempt to relate such measures to CCE values.

The availability of an objective way of quantifying tool extension is in itself useful across a wide range of areas, but our study also leaves many research questions open. First, we are curious about how the CCE relates to other measures of experience, in particular, those on body ownership [e.g., 14, 29] and subjective satisfaction [e.g., 10]. We are particularly curious about the relation between the CCE and novelty effects because this measure could be one way to evaluate interfaces objectively where expectations and novelty play less of a role compared to subjective reports.

Second, because our method allows intertwining with interactive tasks, we may for the first time explore the relation between performance and tool extension. For instance, using a Fitts’s law task, it would be possible to vary task difficulty and thereby performance and quantify its influence on tool extension. Earlier work in psychology and neuroscience has not been able to do this.

Third, investigating the methodology with other interfaces seems interesting. We are interested in using it on tangible user interfaces, possibly comparing it to the approach by Alzayat and colleagues [1]. The purpose for using VR in this work was to examine whether our measure of tool extension applies to free-hand interaction. Virtual tools were not used, because those could bring other effects in play, such as a conflict between a lack of haptic feedback from not grasping a tool but seeing an avatar hand grasping a virtual one. Using the measure with VR controllers or other devices which give tactile feedback also needs careful validation because of the possible sensory attenuation or even interference with the tactile stimuli during the cross-modal congruency task.

Fourth, investigating the influence of distances on tool extension provided by computer-based tools seem particularly relevant for free-hand interaction. The distance in which tool extension applies has remained unclear as manual tools used in neuropsychology pose physical limitations to their users, such as a weight that increases with their length. Computer-based tools do not pose such limitations. For example, Feuchtner and Müller [14] presented elongated and disconnected augmented reality hands to act beyond user’s peripersonal space, and their modulation on the experienced ownership. We kept distance to targets constant in Experiment 1; distance from the hand (physically or in VR) to the target likely influences extension of the body schema but we leave this question for future work.

The measure we have presented is easy to use but requires experimenters to get many details correct (as we experienced in numerous pilot studies). In particular, instructions to participants are crucial, both concerning which stimuli they need to react to (the tactile ones, not the visual ones) and that they need to fixate at the targets for visual distractors (rather than elsewhere or closing their eyes).

7 CONCLUSION
Tool extension suggests that using tools may extend our representation of action possibilities in our surroundings. It is widely studied in psychology and neuroscience but so far only discussed in human-computer interaction. We have introduced a measure of tool extension to HCI, based on the visual-tactile interference paradigm. In two experiments we have shown how this measure can discriminate through interfaces and produce an index of how much they are sensed as extensions of one’s body. This methodology helps to study embodiment and represents a significant advance in working quantitatively with the incorporation of computer-based tools in our peripersonal space.

8 ACKNOWLEDGEMENTS
This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement 648785). We thank Antti Oulasvirta and Adrian Alsmith for their helpful comments.

REFERENCES


