On mobile devices, vibrotactile messages are a common way to give feedback to the user. They might be a less obtrusive means to communicate information about the system status compared to auditory feedback. Much research has focused on the possibilities to perceive and discriminate different vibrotactile messages, less regarding her contextual interpretation. We describe a series of two studies. Aim of the pilot study was to find meaningful vibrotactile messages of which we then wanted to investigate the affective impression and functional connotation on a mobile device within varying staged contexts. Results show that the affective impression of those so-called Tactons is independent of the context. Moreover, we observed a relation between ratings of affective quality and functional applicability. We conclude that tactile feedback messages are unobtrusive, but have to be designed carefully to convey their intended meaning in a working context as well as in a leisure time situation.

Tactile feedback. Mobile context. Affective impression.

1. INTRODUCTION

Modern mobile devices like smartphones and tablets are commonly operated through a graphical touchscreen interface, which needs artificially-generated feedback to acknowledge user actions. As these devices usually have an integrated vibration motor, vibrotactile feedback is getting more and more important to complement or substitute auditory and visual messages. The importance of haptics as an output modality in HCI is shown by the first standard documentary for haptic interaction published by ISO some years ago (ISO 2009), and by ongoing work of expert groups in this field (Van Erp et al. 2010). The ISO standard contains guidelines on many different aspects on the whole topic of tactile and haptic interaction, but does not consider the quality of tactile feedback itself, and particularly not emotional consequences of the feedback, similar to most of the research on tactile interaction.

One way to give tactile feedback on mobile devices are so-called Tactons. These are "structured, abstract messages that can be used to communicate messages non-visually" (Brewster and Brown 2004). The possibilities of those Tactons in the context of mobile touchscreen devices have been investigated intensively by Brown and colleagues (Brown et al. 2006, Brown and Kaaresoja 2006).

Regarding input performance, it could be shown that pen input augmented with tactile feedback can result in higher user performance in a drawing task (Poupyrev et al. 2004). Similar results have been found for finger-based input (Brewster et al. 2007).

Hoggan and colleagues compared devices with a physical keyboard, a standard touchscreen, and a touchscreen with tactile feedback in both static and mobile environments. They found that the addition of tactile feedback to a touchscreen significantly improved finger-based text entry, bringing it close to the performance of a real physical keyboard (Hoggan et al. 2008).

Other researchers focused on the pleasantness of tactile feedback (Koskinen et al. 2008) comparing two ways of generating tactile feedback - piezo actuators and a standard vibration motor; their results also support the advantage of tactile feedback compared to no feedback independent of the technology used, even if their findings suggest piezo actuated feedback to be slightly more pleasant.

The question on distinguishing different vibration patterns was investigated by Feige (2009), who used a remotely controlled wristband to generate different Tactons. He found that five rhythms were easily distinguishable by most of the participants without a difference between the walking experimental group and the sitting control group.

A complementary novel approach (Altinsoy and Merchel 2012) uses an electrotactile display integrated into touchscreen devices to simulate different textures. First experiments focused on the possibility to simulate perceived roughness via electrotactile feedback. They showed that electrotactile stimuli with a high current magnitude
and a low pulse frequency might be able to represent rough surfaces and could provide an additional means to generate tactile feedback messages. Going beyond unimodal tactile feedback, also studies with crossmodal feedback have been done. Hoggan (2010) found out that crossmodal audio-tactile icons are helpful in different usage situations of mobile touchscreen devices, and derived a number of guidelines on the design and application of crossmodal feedback.

Affective feedback has only rarely been the object of investigation (exceptions have been cited above under the term “pleasantness”). However, as these feedback messages usually are not in the users' focus, but are processed along the side up to unconsciously, their first reception concerns mostly their basic emotional connotation, i.e. “bad or good”, and with little conscious reflection (Lang, 2010). Therefore, besides the functional adequacy also affective aspects are important for the design of feedback. One recent study, which is only remotely connected to mobile HCI, focuses on affective reactions on steering wheel vibrations in a car (Äjovalasit et al., forthcoming). They found a correlation between the affective dimensions of arousal and valence and vibration intensity as well as other properties (e.g. the presence of high peak events or high frequency band amplitudes).

For auditory feedback, there has been more research on the relation to emotions. Nevertheless, also in this area most studies with emphasis on emotions focus on in-car situations (e.g. Larsson et al. 2009).

A study by the authors of this paper tried to include both functionality and emotional impact of auditory feedback. We found a context effect for the affective impression of real-life sounds as well as for designed earcons (Seebode et al. 2012). This study setup has been extended to the tactile modality with the context study described in this paper.

The remainder of this paper will present a series of two studies to investigate the relationship between functional applicability and affective impression of vibrotactile feedback on mobile devices. Section 2 describes a pilot study that served as a first step to identify appropriate vibration patterns that are clearly distinguishable to elicit associations to functions that occur in the interaction with mobile devices. Based on these results, a set of 16 Tactons was used for the context study described in Section 3. In this context study the Tactons were rated by participants in order to examine their functional applicability and affective impression as well as to gain insight into possible context effects.

2. PILOT STUDY

In a first pilot study the aim was to find vibrotactile messages that are distinguishable and possibly meaningful for users.

2.1 Participants and Material

18 participants aged between 26 and 37 (M = 30.94, SD = 3.12), five of them female, took part in the pilot study. They were all researchers in computer science, human factors, psychology and related disciplines. All of them had extensive experience with personal computers and mobile phones (100 % PC usage daily and since more than 10 years, 88.9 % smartphone users).

We used the Google developer phone Nexus One with its in-built vibration motor to present the Tactons.

30 different Tactons were designed based on literature review and informal pre-tests. This set of vibrotactile messages contained very short patterns but also more complex rhythms. 14 of the messages were simple patterns produced with the Android API1 the other 16 messages were created using the Immersion Universal Haptic Layer (UHL)2. For reasons of clarity, the description in Table 1 is limited to the subset of 16 Tactons used in the context study.

Figure 1 shows the interface to test the different vibration patterns. There were four consecutive screens with a number of eight (resp. six on the last screen) numbered buttons to replay the different patterns presented on each screen. It was designed the way that always two similar patterns were in one row and had to be compared with each other first, before focusing on the next row of two examples.

Figure 1: Interface to replay the different Tactons in the pilot study.

The simple vibration patterns were created varying only the times for which to turn the vibration motor on and off.

---

2 http://www2.immersion.com/developers/
No further possible features can be influenced directly with the Android API. However, due to the inertia of the in-built motor the duration was not always sufficient for the motor to oscillate with its target amplitude. Therefore, shorter vibrations have lower intensities and so should be perceived as less powerful until a certain threshold.

Figure 2 shows the recorded wave forms (produced by the oscillating mobile device positioned directly on a boundary microphone) of four example vibrations. The first one is a simple short vibration specified with 100 ms; the other three vibrations are examples of the patterns used in the context study. It is obvious that the intensity of the simple 100 ms vibration is considerably lower than of the 300 ms vibration (Pattern 4) when looking at the amplitude. Furthermore, it can be seen that the 100 ms interval between 2 vibrations of 100 ms each is too short to really stop the motor and therefore both oscillations interfere resulting in an even lower intensity (Pattern 2). Additionally, Figure 2 clearly shows that the specified durations given in the code do not fully comply with the real measured durations. This is again due to the inertia of the vibration motor.

The fourth example is a very slight vibration created using the UHL (Pattern 14). With this library it is possible to also have vibrations with less power and thus lower intensity. In this case the pattern called “ENGINE2_33” is illustrated which is a relatively constant oscillation with only 33 % power invoked for three seconds.

2.2 Procedure

First, the participants were introduced to the concept of Tactons and the purpose of this test. As the pilot study was an internal review, they were instructed to think aloud while testing the vibration patterns; accordingly we used a protocol to take notes of the qualitative and quantitative answers instead of a standardized questionnaire. In order to keep the test condition constant for all vibrations, participants were asked to hold the phone in their non-dominant hand and keep it in the same position during the whole test session. The task was to start the two patterns in one row consecutively and answer the given questions, before focusing on the next two vibrations. They were asked three general questions:

- **Discriminability**: Are the two patterns in one row easily distinguishable from each other (and different from the other vibrations)?
- **General impression** (derived from previous studies on Tacton design (Brown et al. 2006, Brown and Kaaresoja 2006)): How would you describe each pattern (in terms of e.g. strength, length, pleasantness and possible associations)?
- **Functionality**: What event of an interactive system on your smartphone would you expect with this vibration pattern (e.g. a started, running or stopped process, a system question, an error or something else)?

They always had the possibility to give additional comments or to further explain their ideas. Moreover, they could also compare patterns from different rows, if they felt that there were similarities. All answers were written down by the experimenter.

To see whether the sound that originated from the motor producing the vibrations had a significant effect, half of the participants were asked to wear noise-cancelling head phones during the test.
2.3 Results

As one aim was to find out, which similar Tactons are distinguishable, a first step in the analysis was to select all patterns that were perceived as unique or well distinguishable by most of the participants. This was the case for 14 of the 30 patterns that were rated as distinguishable by at least 16 of the participants. These 14 Tactons were taken into account for further analysis.

The second step focused on the pairs of patterns that were rated to be very similar. For these remaining eight pairs of patterns (i.e. 16 single Tactons) that were perceived as indistinguishable, each one that had more comments or was rated slightly more pleasant in direct comparison was selected for further examination.

The last step of the analysis focused on the comments and different functionality indications of the remaining 22 patterns. All patterns that were perceived as too complex or only applicable for e.g. incoming calls (N = 4) or no function at all (N = 6) were eliminated from the test set for the context study. Accordingly, 12 Tactons were named as suitable for similar functions by at least ten subjects and were therefore chosen for the context study.

Additionally, new variations of these selected 12 patterns (in terms of repetitions or length) were included in the user study to test whether their quality would change due to these variations and because of comments by the participants like

“This would be a good vibration for [event x] if it was repeated.”

or

“This pattern could be annoying if it was any longer”.

Hence, a resulting set of 16 patterns was employed in the user study. They are listed in Table 1; Patterns 3, 6, 15 and 16 resulted from the before mentioned variations of the pilot study patterns.

The sound of the motor did not have an influence on the perception of uniqueness (chi² = 0.89, p = 0.35). Though, for the possible functions (chi² = 22.80, p = 0.007) significant differences were observed between the participants wearing headphone and the participants not wearing headphones; namely in the group not wearing headphones, “no suitable function” was mentioned more often.

3. CONTEXT STUDY

In every-day usage of mobile devices, feedback messages are given under varying conditions. To mimic different usage scenarios, the following context study focuses on their assessment in an applied setting with three different context situations. The study setup has been applied to sounds for mobile devices in a previous study by the same authors (Seebode et al. 2012), where a context effect for the affective impression of different sounds could be found.

3.1 Participants and Material

Participants in this study were 36 students aged between 20 and 35 (M = 26.56, SD = 3.32), 17 of them female. All of them had extensive prior experience with personal computers and mobile phones, but there were fewer smartphone users than in the pilot study (97.2 % daily PC usage, 69.4 % smartphone users). Again the Google developer phone Nexus One was used to present the Tactons and additionally, to collect user ratings.

The presented Tactons were 16 short vibrations selected after the pilot study (Table 1). The UHL vibrations are shortly described by the Immersion Cooperation (2013).

Table 1: Names and descriptions of the 16 Tactons used in the context study.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>UHL name / pattern</th>
<th># of repetitions</th>
<th>duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>0</td>
<td>40 ms</td>
</tr>
<tr>
<td>2</td>
<td>100 100</td>
<td>1</td>
<td>400 ms</td>
</tr>
<tr>
<td>3</td>
<td>100 100</td>
<td>3</td>
<td>800 ms</td>
</tr>
<tr>
<td>4</td>
<td>300</td>
<td>0</td>
<td>300 ms</td>
</tr>
<tr>
<td>5</td>
<td>300 100</td>
<td>3</td>
<td>1200 ms</td>
</tr>
<tr>
<td>6</td>
<td>300 100</td>
<td>13</td>
<td>5200 ms</td>
</tr>
<tr>
<td>7</td>
<td>SHORT_BUZZ_100</td>
<td>0</td>
<td>400 ms</td>
</tr>
<tr>
<td>8</td>
<td>ALERT2</td>
<td>0</td>
<td>3000 ms</td>
</tr>
<tr>
<td>9</td>
<td>ALERT9</td>
<td>0</td>
<td>3000 ms</td>
</tr>
<tr>
<td>10</td>
<td>EXPLOSION1</td>
<td>0</td>
<td>600 ms</td>
</tr>
<tr>
<td>11</td>
<td>TRANSITION_BUMP_100</td>
<td>0</td>
<td>700 ms</td>
</tr>
<tr>
<td>12</td>
<td>LONG_TRANSITION_RAMP_DOWN_100</td>
<td>0</td>
<td>1000 ms</td>
</tr>
<tr>
<td>13</td>
<td>SHORT_TRANSITION_RAMP_UP_66</td>
<td>0</td>
<td>400 ms</td>
</tr>
<tr>
<td>14</td>
<td>ENGINE2_33</td>
<td>0</td>
<td>3000 ms</td>
</tr>
<tr>
<td>15</td>
<td>ENGINE2_33</td>
<td>0</td>
<td>13000 ms</td>
</tr>
<tr>
<td>16</td>
<td>ENGINE2_33</td>
<td>0</td>
<td>6000 ms</td>
</tr>
</tbody>
</table>

The pattern description for the simple patterns here is abbreviated with the number of repetitions in the next column. In the android code e.g. Pattern 2 would be given: 0 100 100 100 100 meaning, that it starts immediately (0), vibrates for 100 ms, than stops the motor for 100 ms and turns it on again for 100 ms than stops again for 100 ms. Correspondingly, the other simple patterns are described here with the number of repetitions, which leads to the summed durations. The UHL created patterns are described here with their name in upper case letters.

As duration here the specified durations given in the code are labelled, that do not always fully comply with the real measured durations as explained in Section 2.1.
The affective impression of all presented Tactons was measured with the Self-Assessment Manikin (SAM). The SAM questionnaire is theoretically grounded in the dimensional theory of emotions and was developed by Bradley and Lang (1994). It has previously been used for visual stimuli (Lang et al. 2005) as well as for acoustic material (Bradley and Lang 2007). The affective impression of the stimuli is measured on the three affective dimensions of valence, arousal, and dominance. Each is represented by a pictogram-based scale.

Valence is ranging from a happily smiling face for a pleasant stimulus to an unhappily frowning face for an unpleasant stimulus. Arousal varies between an excited figure with eyes wide open (aroused) and a relaxed and sleepy figure (calm). For dominance, ranging between being in control and being controlled, a manikin of increasing size is shown: for high dominance values the manikin covers almost the complete frame.

The original instruction (Bradley and Lang 1994) was adapted and translated to German. Due to the delimited space on the mobile device, we reduced the original nine step scale to a five step scale. While in the original version ratings between the pictograms are possible, we used only the pictograms themselves as points on the scale.

To assess the functional applicability of the Tactons, participants were asked to rate their appropriateness for a given list of six events. For that purpose, we used a five step scale varying between doesn't fit at all and fits very well. The list of events contained the actions of a started, running and stopped process, a system question, system confirmation and an error/warning. The events were chosen based on an expert review and subsequent discussion.

Additionally, we were interested in differences between the experiment blocks according the mental effort. To measure the perceived effort of each context condition, we used the SEA scale (Subjektiv Erlebte Anstrengung), a unipolar instrument ranging between 0 and 220 with higher values indicating higher effort (Ellers et al. 1986). The SEA scale is the German version of the SMEQ also known as RSME (Zijlstra 1985, Sauro and Dumas 2009). We chose the SEA as it is a lightweight instrument shown to have excellent psychometric properties (Sauro and Dumas 2009) even in comparison to more elaborate measures (De Waard 1966).

### 3.2 Context Conditions

As we wanted to collect ratings for each vibration separately, we designed three typical contexts where feedback from a mobile device might occur. Figure 3 shows the setup for our three context conditions.

As a baseline context, common for perceptual experiments, we used a neutral laboratory setup in a sound-isolated test cabinet without any disturbing material. The other two conditions were an office and a bar surrounding. Those contexts were selected as they are relatively easy to set up in a laboratory, and are also typical real-live scenarios.

For the office context we used a cubicle in an open-space office with one desk and some working material as decoration. Adequate background sounds were replayed from four loudspeakers out of a loudspeaker array above the participants’ heads (see Figure 3 middle). For the bar condition the same room was redecorated with coloured light, bar equipment and a barstool to sit on. Again, suitable background sounds and additional music on a low volume ($\text{Leq}_{\text{Office}} = 39.5\, \text{dB}(A), \text{Leq}_{\text{Bar}} = 48.5\, \text{dB}(A)$, measured over 10 minutes, equivalent to the experiment block duration) were replayed during the whole experimental block.
3.3 Procedure

Each session lasted around 60 minutes. The experiment was divided into three blocks, one for each context condition. The blocks were counterbalanced through all participants. At first, the participants were asked to fill in a short introductory questionnaire regarding demographic information and information about their experience with computers, mobile phones, and touch screen systems.

Next, we presented a short training trial. The participants were instructed to rate one short sample vibration (100 ms duration) on the SAM and according their functional applicability.

Before the first block of the experiment started, a short video (about 90 seconds long) of a train journey was shown to the participants. This video was shown before every block, in order to adjust the mood of the participants to a fairly neutral state and to separate the different conditions.

Then, the first block started. The participants were seated in the room which was either decorated as an office or as a bar, or in the neutral test cabinet. They were presented with one Tacton after another; SAM and functionality ratings for each Tacton were collected directly thereafter. The Tactons could be repeated as often as the participants wanted to, but they were asked to rate nevertheless spontaneously. In order to ensure that the vibrations were the same in all three blocks, the participants were instructed to hold the phone in their non-dominant hand and try to keep it the same position during the whole experiment. After each block, the participants filled in the SAM again. They were asked to rate how they felt during the according experiment block overall. In addition, ratings regarding the perceived mental effort of the last experimental block were collected with the SEA scale. This procedure was repeated for the second and third context accordingly.

Afterwards, the participants were asked how realistic the different scenarios were implemented. Answers were collected on a five step scale ranging from not realistic at all to very realistic. Finally, subjects were paid 10€ as compensation for the session.

3.4 Results

This section describes the results on the collected ratings on the affective impression of our Tactons first. Afterwards, results according the functional applicability are presented, following general results in the end.

3.4.1. Affective Impression

Figure 4 shows mean valence and arousal ratings of all 16 patterns separated for the three different context conditions. As dominance ratings did not vary substantially across stimuli. Hence, we focus on valence and arousal here. It can be seen that valence and arousal ratings are negatively correlated (p < 0.01, Pearson’s r = -0.49). Additionally, there is no systematic difference between the three contexts but the ratings vary for the different Tactons. This first impression is supported by the results of a multivariate analysis.
of variance (MANOVA) for the SAM variables with context and pattern as repeated-measured factors: There is no significant difference for context (F(context, 6, 3358) = 1.48, p = 0.18, part. eta² = 0.003) but for pattern (F(pattern, 45, 5040) = 13.97, p = 0.00, part. eta² = 0.11). An interaction effect for context*pattern was not observed (F(90, 5040) = 0.64, p = 0.997, part. eta² = 0.01).

The most pleasant and least arousing stimulus is the very short and slight Pattern 1. Bonferroni-corrected post-hoc tests show that Pattern 1 is rated significantly less arousing than all the other patterns (M diff max = -2.35, p < 0.01; M diff min = -0.58, p < 0.01) and more pleasant than most of the other patterns (M diff max = 2.06, p < 0.01; M diff min = 0.21, p = 1.00).

The opposite example of a very arousing and unpleasant Tacton is Pattern 6 with a very high number of 13 fast repetitions and a comparably long duration of each single vibration. This pattern is also rated more arousing (M diff max = 2.35, p < 0.01; M diff min = 0.66, p < 0.01) and less pleasant (M diff max = -2.06, p < 0.01; M diff min = -0.69, p < 0.01) than all other patterns.

3.4.2 Functional Connotation

Table 2 shows a colour map for the functionality ratings (averaged over the three context conditions) for all 16 patterns. Again the ratings vary across the scale and between the different Tactons.

As expected, the short and unobtrusive Tactons are rated most suitable for positively connoted events that are not critical (e.g. Pattern 1 and 13). Accordingly, stronger and more arousing vibrations like Pattern 5 and Pattern 6 are rated most suitable for error messages. Furthermore, the last three patterns (14 to 16) with low intensity but very long duration are rated most suitable for a running process but not for all the other functions.

A second MANOVA for the functionality ratings again with context and pattern as factors reveals a slight context effect but with a very low effect size (F(context, 12, 3352) = 2.00, p < 0.01, part. eta² = 0.007) and significant differences between the patterns (F(pattern, 90, 10080) = 17.95, p < 0.01, part. eta² = 0.14), but again no interaction of context*pattern (F(180, 10080) = 0.77, p = 0.988, part. eta² = 0.01). Bonferroni-corrected post-hoc tests for the factor context show significant differences between the neutral and the bar setting for the event of a started process (M diff = -0.21, p < 0.01) and between the neutral and the office condition for a system question (M diff = -0.20, p < 0.01) and a system information (M diff = -0.19, p = 0.012). These differences do not reveal any systematic effect of context but the three functions have a positively connoted intention in common.

<table>
<thead>
<tr>
<th>Pattern #</th>
<th>Start</th>
<th>Run</th>
<th>Stop</th>
<th>System Question</th>
<th>System Information</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern 1</td>
<td>0.62</td>
<td>-0.62</td>
<td>0.26</td>
<td>-0.29</td>
<td>0.60</td>
<td>-1.37</td>
</tr>
<tr>
<td>Pattern 2</td>
<td>0.58</td>
<td>-0.73</td>
<td>0.79</td>
<td>0.71</td>
<td>0.88</td>
<td>0.08</td>
</tr>
<tr>
<td>Pattern 3</td>
<td>-0.28</td>
<td>-0.50</td>
<td>0.19</td>
<td>0.19</td>
<td>0.14</td>
<td>0.42</td>
</tr>
<tr>
<td>Pattern 4</td>
<td>0.68</td>
<td>-0.76</td>
<td>0.69</td>
<td>0.74</td>
<td>0.63</td>
<td>0.40</td>
</tr>
<tr>
<td>Pattern 5</td>
<td>-0.59</td>
<td>-0.71</td>
<td>-0.24</td>
<td>0.01</td>
<td>-0.13</td>
<td>0.67</td>
</tr>
<tr>
<td>Pattern 6</td>
<td>-1.46</td>
<td>-0.81</td>
<td>-1.14</td>
<td>-0.82</td>
<td>-0.96</td>
<td>0.63</td>
</tr>
<tr>
<td>Pattern 7</td>
<td>0.80</td>
<td>-0.65</td>
<td>0.82</td>
<td>0.81</td>
<td>0.85</td>
<td>0.44</td>
</tr>
<tr>
<td>Pattern 8</td>
<td>-0.78</td>
<td>-0.42</td>
<td>-0.43</td>
<td>0.00</td>
<td>-0.19</td>
<td>0.27</td>
</tr>
<tr>
<td>Pattern 9</td>
<td>-0.32</td>
<td>-0.27</td>
<td>-0.06</td>
<td>0.18</td>
<td>0.18</td>
<td>-0.14</td>
</tr>
<tr>
<td>Pattern 10</td>
<td>-0.17</td>
<td>-0.67</td>
<td>0.26</td>
<td>0.24</td>
<td>0.23</td>
<td>0.33</td>
</tr>
<tr>
<td>Pattern 11</td>
<td>0.21</td>
<td>-0.65</td>
<td>0.51</td>
<td>0.01</td>
<td>0.41</td>
<td>-0.69</td>
</tr>
<tr>
<td>Pattern 12</td>
<td>-0.45</td>
<td>-0.66</td>
<td>0.24</td>
<td>0.05</td>
<td>-0.20</td>
<td>0.33</td>
</tr>
<tr>
<td>Pattern 13</td>
<td>0.56</td>
<td>-0.58</td>
<td>0.36</td>
<td>0.28</td>
<td>0.53</td>
<td>-0.65</td>
</tr>
<tr>
<td>Pattern 14</td>
<td>-0.85</td>
<td>0.55</td>
<td>-1.08</td>
<td>-1.13</td>
<td>-1.03</td>
<td>-1.16</td>
</tr>
<tr>
<td>Pattern 15</td>
<td>-1.52</td>
<td>0.23</td>
<td>-1.57</td>
<td>-1.48</td>
<td>-1.40</td>
<td>-1.09</td>
</tr>
<tr>
<td>Pattern 16</td>
<td>-1.19</td>
<td>0.49</td>
<td>-1.41</td>
<td>-1.24</td>
<td>-1.20</td>
<td>-1.04</td>
</tr>
</tbody>
</table>
3.4.3 General Results

Correlations between SAM ratings and the event appropriateness ratings can be found. Valence is positively correlated with all events (p< = 0.01, Pearson’s r_max = 0.43, r_min = 0.17) except errors (p< = 0.01, r = -0.11), and arousal shows negative correlations with all events (p< = 0.01, r_max = -0.27, r_min = -0.11) except errors (p< = 0.01, r = 0.31) but not with the event of a system question (p = 0.17, r = -0.33). This is nicely shown by Pattern 11 and Pattern 13. Both are rated most appropriate for the positively connotated functions like an ended process or system information but not for an error message or (presumably due to their short duration) a running process and are rated very pleasant and calm.

According to the life-like implementation of the three contexts, all are rated fairly realistic (M_office=0.94, M_bar=1.06, M_neutral=1.75) on the scale from -2 for “not realistic at all” to 2 for “very realistic”.

Table 3 shows mean ratings on the three SAM scales and the SEA-scale, collected after each experimental block. It can be seen that the neutral context is rated most pleasant and least arousing with the lowest effort. As there were no disturbing material and no background noise in this setting, these findings were expectable.

A third MANOVA with the factor context reveals significant differences for all the SAM dimensions as well as for the perceived effort (F_context(8, 136) = 6.86, p = < 0.01, part. eta² = 0.29). Bonferroni-corrected post-hoc tests show significant differences between the neutral and the office setting on the valence and dominance scale (M_diff valence = 0.58, M_diff dominance = 0.53, p < 0.01). Regarding the arousal ratings, differences are observed between the neutral and the bar context (M_diff = -0.13, p < 0.01) as well as between the neutral and office setting (M_diff = -0.97, p < 0.01).

Table 3: Mean ratings and standard deviations of the three Sam scales and the SEA-scale for each context.

<table>
<thead>
<tr>
<th>Context</th>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>SAM arousal</td>
<td>1.67</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>SAM valence</td>
<td>4.25</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>SAM dominance</td>
<td>3.78</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>SEA</td>
<td>34.22</td>
<td>25.78</td>
</tr>
<tr>
<td>Bar</td>
<td>SAM arousal</td>
<td>2.94</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>SAM valence</td>
<td>3.97</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>SAM dominance</td>
<td>3.56</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>SEA</td>
<td>57.11</td>
<td>41.38</td>
</tr>
<tr>
<td>Office</td>
<td>SAM arousal</td>
<td>2.64</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>SAM valence</td>
<td>3.67</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>SAM dominance</td>
<td>3.25</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>SEA</td>
<td>58.56</td>
<td>41.30</td>
</tr>
</tbody>
</table>

The ratings of mental effort vary significantly between the neutral and the bar context (M_diff = -22.89, p < 0.01) as well as between the neutral and office setting (M_diff = -24.33, p < 0.01). Altogether it can be said that tactile feedback messages, varying on a limited number of dimensions, are able to evoke an affective impression that is not affected by the context situation and that their functional connotation is related to the dimensions of valence and arousal.

4. DISCUSSION

In this paper we report a pilot study followed by a context study on the affective and functional aspects of vibrotactile feedback on mobile devices. The pilot study served as a first step to gather information about simple Tactons on a mobile device. These simple patterns were generated by varying almost only the parameters of duration and rhythm. But due to the physical characteristics of the in-built vibration motor used, different durations are commonly accompanied by changed intensities. We could show that it is possible to create simple tactile feedback that evokes distinct associations and is able to transmit information about the current system state. We are aware that a contentual description of vibrations, beyond technical parameters is not straightforward. Thus, participants in the pilot study were asked to describe the patterns given some example parameters (strength, length, but also pleasantness and possible associations). To exemplify, the patterns depicted in Figure 2 where described as “slight, hollow, and nearly two different frequencies” (Pattern2); “unpleasant, intense, alarming, annoying or aggressive, and attention craving” (Pattern 4); and “like a machine/ motor/ driller, interesting, and not annoying but noticeable” (Pattern 14). Thus, linking meaning directly to waveforms is possible, although it depends on the users’ personal experiences and associations which may lead to varying affective impressions. These concern mostly the dimensions valence and arousal, less dominance, probably because we did not have an interactive setting with an interaction partner present, and thus the idea of controlling someone was less pronounced.

Based on these results, the context study with a sub-set of 16 Tactons examined the effect of context on the affective impression and functional connotation, the rationale being that mobile phone users rely on the vibrotactile feedback especially in situations where the main focus is on something else, be it work or leisure. In these situations, feedback messages usually are not in the users’ focus, but are processed along the side up to unconsciously. Hence, their first reception concerns mostly their basic emotional connotation, i.e. ‘bad or good’, and with little conscious reflection, which
could be influenced by the work or leisure surrounding itself, but also due to the social situation. As we found a context effect for sounds (Seebode et al. 2012), it is interesting that vibrations seem to be unaffected by this manipulation of context aspects. Not surprisingly, in the former study, sounds were more affected by the disturbing noise in the background. Thus, the aspects of context varying in this study might not be the most important ones for tactile messages. Furthermore, the information that those tactile messages contain is less rich and diverse as it is the case for sounds.

However, a relation between vibration features (such as duration and intensity) and affective impression as well as functionality could be found. This complies with findings by Brown et al. (Brown et al. 2005) also suggesting that too strong vibrations can be unpleasant, while too weak vibrations may be difficult to be perceived. Accordingly, using intensity in terms of amplitude as a sole parameter for Tacton design can be problematic. Stimuli with short duration (less than one second) and thus lower intensity are rated calm and pleasant. If these slight patterns were considerably longer, their pleasantness decreased with increasing arousal like the three Patterns 14, 15 and 16 with Pattern 15 as the longest one in the whole Tacton set used here. However, Pattern 6 was rated significantly more arousing than Pattern 15 and 16 even though its duration is lower. This shows that rhythm is one very important feature, following the findings of earlier research (Hoggan 2010).

Unsurprisingly, it is possible to create annoying tactile feedback messages on the mobile device if they exceed a certain duration and intensity. Due to the physical nature of the inertial mass generating the vibration, increases in intensity can be caused by an increase in duration alone, which is probably less obvious. Hence, system developers should keep this in mind and be careful when designing tactile feedback to not attenuate the general advantage of this modality of being less bothersome than sounds by creating too complex or intrusive stimuli. Subtle variations are sufficient to design distinguishable Tactons, even though the risk of confusion is higher for tactile feedback than for auditory feedback.

Of course, our findings so far only hold for the vibration patterns used in this experiment with this specific mobile device, and it is plausible, that the impact of other Tactons generated with different hardware might be more affected by context. Nevertheless, we present an easy and effective setup to test this assumption.

5. CONCLUSION

This paper describes two studies investigating vibrotactile feedback on mobile devices. In the first pilot study it was shown that current mobile devices offer the possibility to create tactile feedback that evokes distinct associations and is able to transmit information about the current system state.

Based on the results of the pilot study, we selected a set of different Tactons that were shown to be appropriate for mobile device feedback. Accordingly, the context study focused on the affective impression and functional connotation of these Tactons in three different context conditions, namely a bar, office and a neutral laboratory context. The results show no significant effect of context on the affective impression of our Tactons. Nevertheless, an effect for the different patterns could be shown: the affective impression varied with rhythm as the main feature for pattern difference. Related to their affective impression, also the functional connotation varied across different vibration patterns and met the intended functionality connotations.

It is clear that more demanding contexts – e.g. walking on the street or a dual-task scenario – may have a stronger impact on affect and workload. Hence, we intend to conduct interaction experiments with system prototypes that focus on these aspects in the future. Usability measures will then complement the affective impression ratings to create a broad knowledge about the perceived quality of system feedback. Moreover, we will focus on combined audio-tactile feedback messages that merge the results of both unimodal studies to design pleasant feedback with a possibly clear and distinct meaning.

6. ACKNOWLEDGMENTS

This work was funded by the German Research Foundation (DFG – 1013 ‘Prospective Design of Human-Technology Interaction’) and supported by Telekom Innovation Laboratories.

7. REFERENCES


