Int. J. Human-Computer Studies 107 (2017) 38–61

Contents lists available at ScienceDirect

International Journal of Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Exploiting haptic facets: Users' sensemaking schemas as a path to design and personalization of experience

Hasti Seifi*, Karon E. MacLean

Department of Computer Science, University of British Columbia, 201-2366 Main Mall, Vancouver, British Columbia, Canada

ARTICLE INFO

Keywords:
Haptic language
Affective vibration annotations
Perceptual dimensions
Individual differences
Data collection methodology

ABSTRACT

Our poor understanding of the connection between haptic effect engineering – using controllable parameters like frequency, amplitude and rhythm – and how sensations are comprehended by end-users hinders effective design. *Haptic facets* (categories of attributes that characterize collection items in different ways) are a way to describe, navigate and analyze the cognitive frameworks by which users make sense of qualitative and affective characteristics of haptic sensations. Embedded in tools, they will provide designers and end-users interested in customization with a road-mapped perceptual and cognitive design space. We previously compiled five haptic facets based on how people describe vibrations: *physical*, *sensory*, *emotional*, *metaphoric*, and *usage examples*.

Here, we report a study in which we deployed these facets to identify underlying dimensions and cross-linkages in participants' perception of a 120-item vibration library. We show that the facets are crosslinked in people's minds, and discuss three scenarios where the facet-based organizational schemes, their linkages and consequent redundancies can support design, evaluation and personalization of expressive vibrotactile effects. Finally, we report between-subject variation (individual differences) and within-subject consistency (reliability) in participants' rating and tagging patterns to inform future progress on haptic evaluation. This facet-based approach is also applicable to other kinds of haptic sensations and even other modalities, and can inform multimodal experience design through a descriptive design language shared between different modalities.

1. Introduction

Despite growing interest in and availability of haptic technology in consumer markets, even its most common manifestation of vibrotactile feedback is still limited in everyday use, generally appearing as a dull, undifferentiated and often annoying buzz. While a dearth of expressive hardware is one obvious cause, there are comparable difficulties in *designing* with even the hardware we already have for both vibrotactile and other haptic display modalities (MacLean et al., 2017).

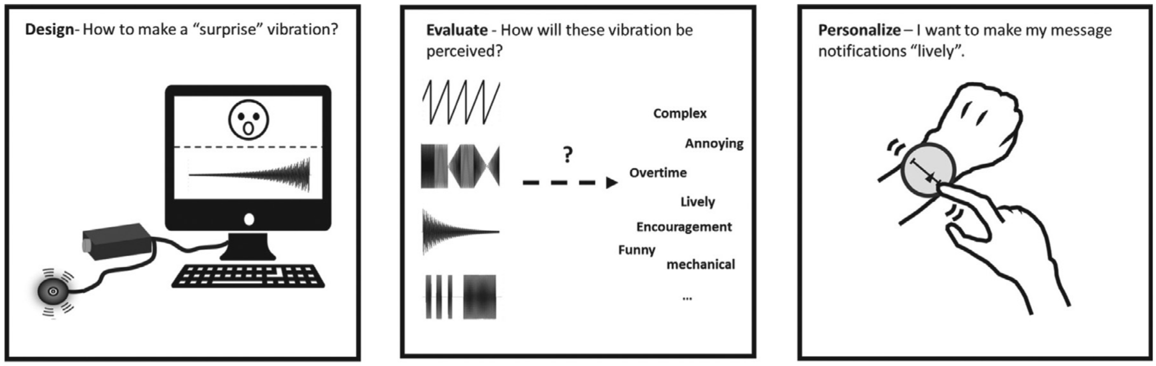
Design is difficult for many reasons, not least due to large variances in individuals' preference and interpretation of how vibrations feel and what they suggest (Lo et al., 1984; Hollins et al., 2000; Peck and Childers, 2003; Levesque et al., 2012; Seifi and MacLean, 2013). Here we highlight two gaps in support which we propose are central.

A Lack of Guidelines and Tools: When making (sketching, refining) and evaluating sensations, designers often identify requirements in terms of usage examples (e.g., allowing presenters to track time during their presentations), intended emotions (sadness, surprise), or accompanying media (a racing car in a game) (Chan et al., 2008; Tam et al., 2013; Carbon and Jakesch, 2013; Zhao et al., 2014; Swindells et al., 2014; Israr et al., 2014), but are forced to *design* with engineering parameters (Scenario 1, Fig. 1a). In other cases, designers have a set of vibrations (whether newly created or accessed within an existing collection) and wish to *evaluate* their aesthetic and qualitative characteristics (Scenario 2, Fig. 1b). The ability to use low-level engineering parameters to construct or evaluate in other idioms is tacit knowledge that haptic designers build over years and through extensive contact with users. It is hard to communicate, incorporate in tools or transfer to others.

Perception is Personal but Personalization is Unsupported: Past studies of vibrotactile applications in real-world contexts indicate the necessity of end-user personalization (Tam et al., 2013; Seifi et al., 2014; Israr et al., 2014). However, there is a dearth even of effective *expert* tools for far more accessible and perceptually understood engineering parameters like vibrotactile amplitude and frequency; easy and practical mechanisms that would make sense to end-users are rare indeed. Unsurprisingly, previous work suggests that personalizing based on engineering parameters is beyond end-user capacity and willingness. However, when given tools in their own language domain, users can quickly access and modify their desired vibrotactile notifications (Scenario 3, Fig. 1c) (Seifi et al., 2015, 2014).

* Corresponding author.
E-mail addresses: seifi@cs.ubc.ca (H. Seifi), maclean@cs.ubc.ca (K.E. MacLean).

http://dx.doi.org/10.1016/j.ijhcs.2017.04.003
Received 2 August 2016; Received in revised form 29 March 2017; Accepted 23 April 2017
Available online 26 April 2017
1071-5819/



(a) **Design Guidelines and Refining**: Designers often need to translate aesthetic requirements specified in emotion, metaphor, and usage spaces (*e.g.*, surprise) to sensory and engineering parameters (*e.g.*, frequency); and to refine candidates.

(b) **Evaluation:** Assessing or accessing the perceptual and aesthetic qualities of vibrations, created by manipulating engineering parameters, allows designers to use them appropriately.

(c) **Personalization:** End-users can more efficiently select and tune vibrations in a perceptual and aesthetic space than in an engineering space, requiring the further capability of *repositioning* sensations within cognitive spaces.

Fig. 1. Three scenarios in vibrotactile design, evaluation, and personalization that facets can support when fully instantiated in design tools.

1.1. Facets: aligning content access with mental frameworks

People unconsciously use a multiplicity of cognitive frameworks or *schemas* to describe qualitative and aesthetic attributes of vibrations (Obrist et al., 2013; Schneider and MacLean, 2014; Seifi et al., 2015). Sometimes people describe a vibration based on its similarity to something they have experienced before (*this is like a cat purring*), on emotions and feelings (*this is boring*), or intended usage (*this tells me to speed up*). These schemas, themselves composed of many attributes (Fig. 2a) are in users' minds: shaped by their past experiences and training, they provide a cognitive scaffolding on which people rely for sense-making.

Facets, a design concept originating from the information retrieval domain (Yee et al., 2003; Smith et al., 2006; Hearst, 2006, 2008; Fagan, 2010), capture the multiplicity and flexibility of users' sense-making schemas for physical and virtual items. A facet encapsulates the properties or labels related to one aspect of or perspective on an item and offers a categorization mechanism. For example, examples of alternative facets for a collection of architectural images are people (such as designer, agency, historical figure), time periods, geographical location (GPS coordinates, province, neighborhood), and structure types (function, architectural elements). For a collection of clothing items they might be garment type (top, bottom, inner, outer, accessories), color, brand, formality, season (Yee et al., 2003; Hearst, 2006). A given facet may be composed of a single property (e.g., brand) or a set of diverse elements that reflect that perspective – e.g., lists of descriptive words (tags), numerical scales, binary or multicategory attributes (e.g., province). The facet characterization varies by domain and relies on a user's knowledge and conceptual mapping of that domain. Multiple facets can be used flexibly together to describe or examine different aspects of a given item in a collection, or alternatively, explore those aspects in light of other collection items.

In a previous work, we identified five facets for vibrations based on the literature (Table 1) which captured: 1) physical$_f$ attributes of vibrations that can be objectively measured such as duration, rhythm structure, etc. 2) sensory$_f$ properties such as roughness, 3) emotional$_f$ connotations, 4) metaphors$_f$ that relate the vibration's feel to familiar examples (e.g., a continuous rampup vibration can be perceived as if it was a speeding car engine), and 5) usage examples$_f$ or events where a vibration fits (e.g., *speed up*).

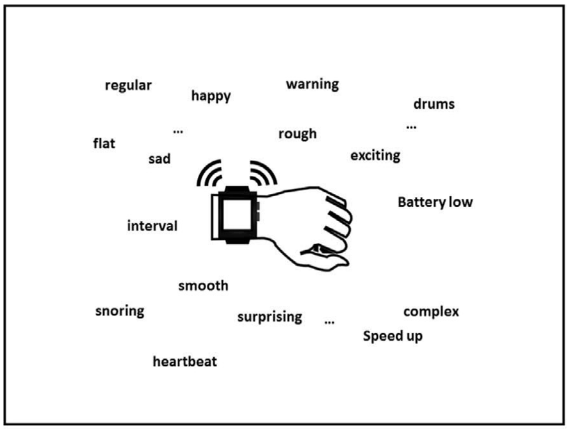
Here, we revise these into four facets: sensation$_f$, emotion$_f$, metaphor$_f$, and usage examples$_f$ (Table 1). For consistency with past haptic literature (Ternes, 2007), we now refer to dimensional attributes that can be objectively measured (e.g. duration, frequency) as *engineering space*. The sensation$_f$ facet now includes the subjective dimensional attributes energy and tempo, previously in the physical facet.

These facets provide unique ways to assign a familiar meaning to a haptic sensation. For example, the metaphor$_f$ and usage$_f$ facets rely on previously experienced sensations and usage contexts to make sense of vibrations (see Seifi et al. (2015) for more details). We implemented these facets in an interactive graphical visualization and navigation tool, *VibViz* (Seifi et al., 2015), and denote them and related concepts here with a special font and subscripts (as explained in Fig. 2).

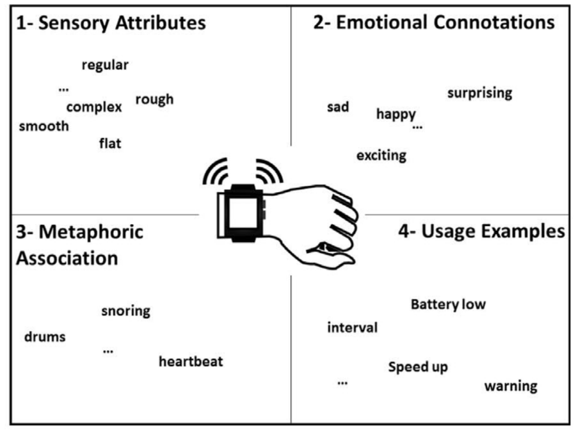
While not meant to be a unique or complete delineation of the possible vibrotactile facet space, this set does provide a practical sense of what facets can offer to design. Because a given vibration can be located in the context of any and all, each highlighting a particular aspect, they can organize a messy hodgepodge of inconsistent language and mixed models into a powerful tool that leverages perception and analogy. The interactive visualization tool *VibViz* (Seifi et al., 2015) allows untrained users to peruse a large vibrotactile collection by viewing items in multiple facets simultaneously and dynamically.

These multi-facet views thereby become rich, layered descriptions which inform design. For example, *VibViz*'s linked facets show how an individual item may have different perceptual near-neighbors and contrasts in the different facets.

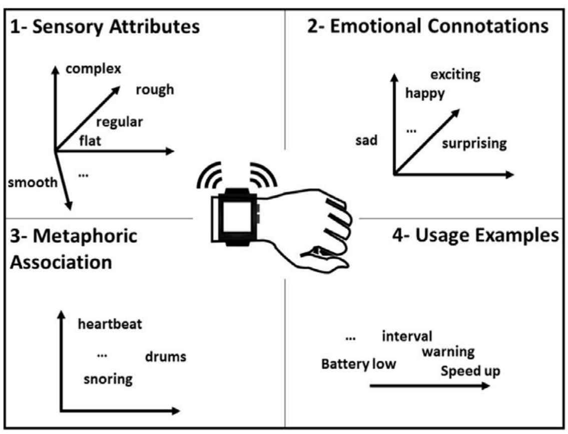
From Browsing to Manipulating in Facet Space: In its primary form, a facet is just a flat list of attributes like tags and ratings (Fig. 2b). Thus, it only allows us to browse existing, defined elements (as *VibViz* does). What if a designer or user wants to *change* an element, or find points in between existing library items (Fig. 1 scenarios)? A semantic *dimension* offers a structure for the facet; it provides a continuous perceptual parameter along which one can move vibrations or characterize them (Fig. 2c). Imagine a slider that makes a vibration more or less "exciting", "alluring" or "bell-like" – in contrast to ones that change its base frequency or amplitude. Such sliders would allow both trained designers and untrained end-users to manipulate (sketch, ideate, personalize) vibrotactile signals more directly by offering handles in a language framework relevant to their purpose.



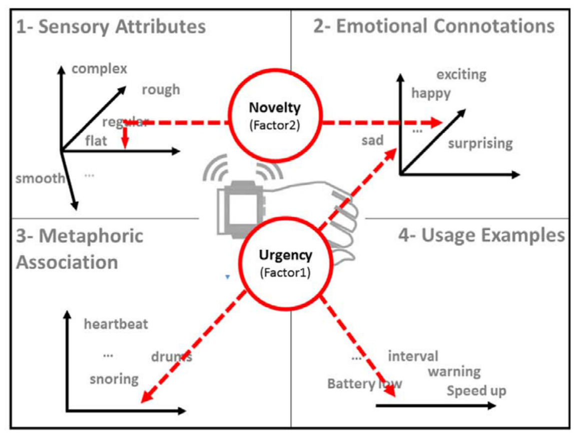
(a) People use ***mixed language*** to describe and make sense of vibrations, which is highly descriptive; but its disorganization makes it hard to use in design.



(b) ***Facets*** organize users' descriptions into *categories of labels*, each describing and orienting elements according to one aspect that the labels in that facet share.



(c) The underlying ***semantic dimensions*** of each facet (shown as black arrows) structures its attributes, and exposes axes along which there is continuity.



(d) Factors are conceptual constructs that can describe the ***linkages between dimensions of the four facets*** (red arrows)).

Fig. 2. Concept sketch showing haptic facets, dimensions and their linkages. Central elements (denoted throughout paper with a special font and subscripts) include (1) tag: a label/word that people use to describe an attribute of a haptic sensation (e.g., soft, exciting); (2) facet$_f$: a framework that binds related attributes of haptic sensations into a descriptive category; (3) dimension$_d$: a continuous parameter that delineates variations in a facet; and (4) factor$_{fact}$: a conceptual construct underlying linkages among different facets (deduced here using factor analysis).

However, to allow continuous movement along cognitively useful dimensions, a tool must do far more than locate discrete sensations within facet space: *it must identify and present a topologically continuous mapping between the facets and engineering spaces, so that every point of the slider's range can be rendered.*

Further, VibViz already hints at considerable redundancy between facets – when a dimension in one facet is very similar to that of another, but goes by a different name. Facets are not independent spaces, but alternative views of the same thing. Mapping connections specifically will enable designers to translate or formulate requirements from one facet space (e.g., emotional or application-driven constraints) into more actionable sensory and engineering spaces (*Scenario 1*, Fig. 1a) or evaluate aesthetic characteristics of a set of vibrations given their sensory properties (*Scenario 2*, Fig. 1b).

1.2. Research questions

A major objective of this research is to establish a means of finding such mappings. As a first step, we have pursued three questions:

(Q1) Within-facet substructure: what are the underlying dimensions of the facets that dominate users' reaction to vibrations? For example, for the emotion$_f$ facet one could then design or identify the most emotionally distinct vibrations. These dimensions are the first step towards perceptually salient continuous "sliders", such as roughness$_d$.

(Q2) Between-facet linkages: how are attributes and dimensions in different facets linked with each other? A specific mapping will allow for translation of requirements from one facet to another (e.g., emotion$_f$ to sensation$_f$ and vice versa) and provide the basis for a topologically continuous mapping between the facet dimensions and engineering parameters. Designing a "surprising" sensation is much simpler if one can access its sensory$_f$ characteristics to be irregular, ramping up, and rough. Our format convention for vibration tags or attributes highlights *points* in a facet space, as opposed to dimensions.

(Q3) Individual differences (I.D.) in facets: to what extent do people coincide or differ in their assessment of vibration attributes? Facets are based on frameworks in users' mind which can vary greatly, for example due to past experiences and culture. Understanding this variation can shed light on individual differences in preferences and meaning-mappings, and inform development of robust haptic evaluation instruments.

Table 1
Vibration facets used here, taken with minor alterations (†) from (Seifi et al., 2015). These facet properties are combinations of ratings (quantitative attributes such as i, ii, iii for sensation$_f$ facet) and tags (list of words iv). For example, in the sensation$_f$ facet, *i*, *ii* and *iii* are single attributes on which an item can be rated, while *iv* is a list of descriptive tag words that might apply to sensations when considered from this viewpoint. *Modifications:* (1) Omitted the physical$_f$ facet. For consistency with past haptic literature, we now refer to dimensional attributes that can be objectively measured (e.g., duration$_d$, frequency$_d$) as *engineering space*. (2) The sensation$_f$ facet now includes the subjective dimensional attributes energy$_d$ and tempo$_d$, previously in the physical$_f$ facet (NounProject, 2016).

Facet	Attributes
1. Sensation$_f$ Perceptual properties of vibration.	i) *energy*† ii) *tempo* or speed† iii) *roughness* iv) *Sensory words: 24 adjectives* from touch dictionary (Guest et al., 2011).
2. Emotion$_f$ Emotional interpretations of vibration.	i) *pleasantness* ii) *arousal* iii) *Emotion words: 26 adjectives* from touch dictionary (Guest et al., 2011).
3. Metaphor$_f$ Familiar examples resembling the vibration's feel.	*Metaphor words:* We collected a set of *45 metaphors* for our list of usage examples, asked colleagues and friends to provide metaphors for our vibrotactile effects, and used the NounProject website (NounProject, 2016) for brainstorming on metaphors.
4. Usage Examples$_f$ Types of events which a vibration fits.	*Usage example words:* We collected and consolidated a set of *24 usage examples* for presentation timing and exercise tracking (Tam et al., 2013).

1.3. Scope

We used the *VibViz* vibration library and the concept of facets to investigate these questions. We first collected an extensive set of user *annotations* (selections of adjective ratings and tags) for library elements to situate the vibrations within the four facets (Seifi et al., 2015). We obtained this data in a two-step process adapted from data collection methods in the music domain (Turnbull et al., 2008), first with three experts and then with 44 lay users.

In our subsequent analysis we derived semantic dimensions of each facet through Multidimensional Scaling (MDS) analysis (Cox and Cox, 2000), and investigated between-facet linkages using factor analysis (Thompson, 2004). With this data, we updated and further populated Table 1's descriptions to include our derived facet dimensions and their linkages. Our analysis occurred at multiple levels: we examined low-level properties and linkages of individual tags (*tag level*), and then semantic facet dimensions obtained from MDS analysis (*dimensional level*), and finally compared these across the four facets (*facet level*). Thus, our novel contributions include:

1. Empirically derived semantic dimensions of four vibrotactile facets;
2. Between-facet linkages at dimensional and individual tag levels, and discussion of their implications for vibrotactile design and tools;
3. Analysis of individual variations in rating and annotating vibrations;
4. A two-step methodology for annotating large sets of vibrotactile effects, and data on its validity and reliability; and
5. A publicly available dataset of 120 vibrations and their annotations and dimensions – the VibViz Dataset (2016).

In the remainder of the paper, we present the related literature on tool development, perceptual dimensions of vibrations, and haptic evaluation methodology (Section 2), and highlight important aspects of our approach (3) followed by data collection details (4) and analysis procedure and results (5). In Section 6, we describe how our results support the design and evaluation scenarios outlined above (Fig. 1) and compare our facet dimensions and linkages to any existing dimensions in the literature. We finish by reviewing our data collection and analysis methodology and presenting interesting directions for future work.

2. Related work

The design process for haptic sensations will inevitably vary substantially depending on designers and use cases, but it usually involves several rounds of design, evaluation, and fine tuning of the stimuli and usage scenarios (Chan et al., 2008; Brunet et al., 2013; Zheng et al., 2013; MacLean et al., 2017). To support this process better, we need effective authoring tools, design knowledge and guidelines, as well as evaluation methodology and metrics. Below we describe progress in these areas and how our work builds on them.

2.1. Tools for vibrotactile design and personalization

With their crucial role in the design process, haptic authoring tools have received increasing attention in the last decade. Design tools by nature facilitate use of some parameters and approaches while limiting access to others; e.g., pre-designed themed color sets vs. full-spectrum palettes – an example of parameter-limiting; or fine tuning and precision vs. rapid prototyping and creative flow, i.e., approach-limiting. Existing haptic tools are built around the most important design parameters and approaches identified in the literature or by practitioners. For example, to support design around rhythm or temporal pattern, the tools facilitate precise modification and referencing of vibrations on a timeline (Ryu and Choi, 2008; Swindells et al., 2014; Schneider and MacLean, 2016). Recent instances promote use of examples and design by demonstration as well as rapid prototyping by allowing easy modification of design parameters (Hong et al., 2013; Schneider et al., 2015; Schneider and MacLean, 2016). However, to our knowledge these tools currently provide access only to low-level engineering parameters. Perceptual and affective controls over vibrations are missing, and this slows design.

Content design and manipulation are no longer done only by a specific group of users (Lieberman et al., 2006). In several other domains (e.g., photo and video editing, music mixing, configuring software), a spectrum of tools exist for various expertise levels (Saul et al., 2011; Evening, 2013; Harrower and Brewer, 2013). Haptic design tools are catching up: while past tools have mostly focused on experts, recent trends have targeted end-user haptic content creation and personalization (Israr et al., 2014; Zhao et al., 2014; Seifi et al., 2014, 2015).

Our work informs design of higher level controls, which can be thought of as tuning sliders or knobs and might be implemented as such in a design interface. These will benefit both expert design tools and end-user personalization interfaces.

2.2. Knowledge of perceptual and qualitative attributes of vibrations

A body of work has investigated perceptual dimensions of natural (e.g., textures) and synthetic haptic stimuli (e.g., vibrations), and users' language for touch (Hollins et al., 2000; van Erp and Spapé, 2003; MacLean and Enriquez, 2003; Chen et al., 2009; Guest et al., 2011; Okamoto et al., 2013; Obrist et al., 2013; Doizaki et al., 2014). In our own previous work, *VibViz*, we compiled five vibrotactile facets based on dimensions and properties known in the literature for vibrations and users' language (Seifi et al., 2015).

Several tactile perceptual studies exist on natural textures (e.g., fabrics, fluids and various surface materials) due to their higher availability and wider range of sensations (see Okamoto et al. (2013) for a survey). However, the resulting dimensions (such as warm/cold, sticky/slippery) are not easily translated to synthetic sensations. Others examine prominent vibrotactile attributes based on users' similarity groupings or ratings for small to large sets of vibrations. They report energy, roughness and rhythm as the most important design parameters (van Erp and Spapé, 2003; Brown et al., 2006; Ternes, 2007; Hoggan and Brewster, 2007a). While these studies give insights into vibration perception, they tend to be organized in terms of engineering or sensory parameters and are not linked to language attributes in users' minds.

Recent studies examine users' tactile language and descriptions as a window onto understanding prominent properties of touch. Notable among these is Guest et al's collection of touch-related English vocabulary (Guest et al., 2011): based on MDS analysis of word similarities, the authors propose three dimensions for sensory words (*roughness*, *dryness* and *warmness*), and three for emotional words (*comfort*, *arousal* and *sensual quality*). We use this collection of sensation and emotion words in our facets; however, the identified dimensions are not validated for synthetic haptic sensations. Further, other aspects of users' languages such as metaphors and usage examples are not examined.

Our own facets were previously constructed based in part on this literature; here, we further confirm, refine and add to these dimensions and attributes by analyzing users' perception of a large library of vibrations collected through the facets.

2.3. Methodology for evaluating qualitative attributes of vibrations

Previous research in related areas typically adapts methodology from other domains for haptic studies, or refines existing haptic evaluation methodology to be more time- and cost-effective. For example, MDS studies in haptics were originally adapted from the auditory domain to investigate perceptual distances between tactile sensations (Grey 1977; Hollins et al., 2000; Cox and Cox, 2000). Other researchers use phenomenology to obtain richer language-based descriptions of haptic sensations (Obrist et al., 2013; Schneider and MacLean, 2014). However, phenomenological studies are time-consuming and thus are only practical with few participants and small sets of sensations. Recently, a project examined the feasibility of using crowdsourcing platforms (e.g., Amazon's Mechanical Turk) for vibration evaluation (Schneider et al., 2016). Despite promising results, the methodology is mainly tested for Likert scale evaluation and is not yet verified for richer, language and annotation-based haptic studies.

Despite some progress in haptic evaluation approaches, it remains singularly difficult for a researcher to collect rich feedback from lay users in a manner that scales to large stimuli sets. Our data collection methodology, adapted from the music domain, by necessity has had to fill this gap. Here, we report its execution details and examine its validity and reliability.

2.4. Instruments for evaluating haptic sensations

As haptic effects are designed for a wide variety of use cases and requirements, researchers frequently must devise a custom evaluation instrument for every study. Recent investigations have laid the foundations for devising a standard yet flexible instrument for vibrations through examining users' language and compiling important vibration properties and common metrics across past studies.

Most relevantly, Guest et al. provide a linguistic instrument for tactile sensations called the "touch perception task" (TPT) (Guest et al., 2011). TPT is composed of 26 sensory ratings and 14 emotional ratings and was tested by its authors on natural textures.

Here, we have re-used the annotation instrument we previously developed for validating and populating *VibViz*, built around language and metrics found in the literature. Specifically, (a) four of our five Likert scale ratings (strength/energy$_d$, roughness$_d$, pleasantness$_d$, and arousal$_d$) are commonly used metrics; while (b) our sensation$_f$ and emotion$_f$ tag lists are based on Guest et al.'s sensation and emotion vocabulary (Guest et al., 2011). We introduced the tempo$_d$ rating scale as well as the metaphor$_f$ and usage example$_f$ tag lists in our previous work on *VibViz* (Seifi et al., 2015). When used to annotate a large vibrotactile library, this more comprehensive annotation instrument can generate results that will inform future vibrotactile evaluation instruments by identifying the redundant facet attributes and providing an estimate of users' reliability and variation in responses.

3. Approach

To investigate the semantic dimensions of these facets and their linkages, we began with *VibViz*'s source vibrations and its comprehensive but efficient evaluation instrument (Seifi et al., 2015). We report the scalable and robust methodology that allowed a comprehensive annotation of our vibration library and use standard dimensionality reduction methods to analyze the resulting dataset. Below, we describe each aspect of our approach in more detail.

3.1. Rich source vibrations

To identify underlying dimensions and linkages of facets, we used a large and varied set of vibrations. In (Seifi et al., 2015), we described our various tools and inspirations including systematically changing vibration parameters (e.g., rhythm, frequency), modifying audio files to serve as vibrations or using audio files as reference for designing vibrations, and running pilot design studies where our lab colleagues designed vibrations for a given set of metaphors (see Seifi et al. (2015) for more details on our library design process). Our design process was intertwined with developing the four facets and their annotation instrument and resulted in 120 vibrations with a wide range of qualitative and affective characteristics.

3.2. Inclusive and concise annotation instrument, for a flat descriptor set

For an accurate picture of the vibrations, we needed an inclusive and non-redundant annotation instrument. If an important rating or tag is not included, we would be unable to identify the corresponding dimension (exclusion risk). In contrast, redundant ratings or tags can introduce noise. As the set of ratings and tags grows, users' (even experts') ability to consistently characterize a vibration decreases (redundancy risk).

We developed our ratings and tags to reduce both risks. We included quantitative rating scales that are frequently utilized in vibrotactile studies, and incorporated as many relevant tags as possible in our evaluation's first step with experts (mitigating exclusion risk). After the expert annotation phase, we removed and consolidated redundant items in a discussion session (mitigating redundancy risk). The ratings capture users' perception of attributes that are previously identified to be salient for vibrations, while the tags allow us to derive other salient dimensions not known before. The list of tags for the sensation$_f$ and emotion$_f$ facets were derived from the literature whereas the metaphor$_f$ and usage$_f$ tags were the result of our pilot data collection with our lab colleagues. We limited the usage context to time tracking contexts (e.g., presentation timing and exercise tracking) as this accounted for a reasonable range of existing vibrotactile applications. Therefore, the list of usage tags was tractable. For the metaphor $_f$ facet, imposing such a limit in a reasonable way was a challenge. As a result, the variety of people's past experiences and metaphors applicable to the vibrations led to a larger number of metaphor tags compared to the other facets. The results of the process are five bipolar 7-point Likert scale ratings, a form frequently used in the literature, and four lists of candidate tags (see Table 1 for an overview, and Appendix A for a full list of tags proposed for each facet).

3.3. Scalable and robust data collection methodology

We needed a comprehensive 'gold truth' annotation set that covered the full *VibViz* library. Ideally, annotations would be applied by individuals who rated the entire facet space for all the items. This would require individuals to rate and tag 120 vibrations, each according to five scales and 121 candidate tags. In piloting, we found this was too mentally and physically demanding to be suitable for lay users with varying levels of commitment, confirmed by poor signal-to-noise properties of that pilot data. We therefore devised a new collection method that could be spread across multiple participants (scalable) and would be robust to outliers, i.e., the occasional low-commitment participant – or at least, to clearly identify these.

Music annotation literature provides interesting alternative approaches for data collection, such as a *panel of experts:* Pandora Internet Radio uses experts to annotate its music dataset, constructing a "gene sequence" for each music piece that is used for music recommendations (Pandora, 2016; Turnbull et al., 2008). Alternatively, services such as Last.fm *crowdsource* the annotation task, incenting end-users to add free-form textual tags to songs from which it derives music "folksonomies" (Jäschke et al., 2007; Turnbull et al., 2008; Last.fm, 2016). However, our access to haptic experts is limited and the literature lacks a set of standard attributes for vibrations. Furthermore, we can not yet fully crowdsource vibration annotations, in large part due to hardware limitations and lack of a validated methodology (Schneider et al., 2016).

We therefore adapted these two approaches into a two-stage evaluation system. In the first *expert annotation* stage, three haptic designers rated and tagged the vibrations employing initial rating scales and tag lists, with encouragement to be liberal in application of tags to stimuli. In the *lay user validation* stage, a larger number of participants with no haptic background adjusted the experts' ratings and tags for subsets of the library – principally by removing tags which they felt did not apply, since this proved to be mentally easier than applying new ones; although tag addition was also allowed. The first stage resulted in consistent annotations across the library that were relatively free of the noise introduced by participants' fatigue and lack of commitment, but reflected only a small number of subjective opinions. In the second stage, we pruned the potentially overpopulated annotation dataset by bringing in additional, but potentially less committed, perspectives. We fully detail the process in Section 4.

This methodology does have a bias risk: participant perceptions of vibrations in the second stage can be influenced by the rating values and tag assignments that they are shown. We devised mechanisms in our experiment design to mitigate this bias, and evaluated its impact on our final dataset.

3.4. Data analysis methods

We used Multidimensional Scaling to identify the underlying dimensions for the tags (but not the rating scales or values) in each facet, and factor analysis to investigate constructs that link dimensions (including rating scale data) between various facets.

Multidimensional Scaling is a dimensionality reduction technique that is commonly used to derive and visualize a low-dimensional perceptual space from a high-dimensional dataset (Cox and Cox, 2000). We used Matlab's classical MDS implementation (a.k.a. Principal Component Analysis or PCA) where the distances among the items (vibrations or tags) are Euclidean – as opposed to ordinal, as in a non-metric MDS (Matlab, 2016).

Factor analysis is typically used to identify underlying variables (a.k.a. factors) that connect and describe a set of observed but correlated quantitative variables (Thompson, 2004; Yong and Pearce, 2013). For example, factor analysis is usually applied to surveys with several Likert-scale questions to find connected questions. We applied factor analysis to our derived facet dimensions, and the ratings collected for our five scales.

4. Data collection and pre-processing

Here, we detail the collection of ratings and tags for the vibrations in two stages described above – expert annotation, and novice validation; then describe dataset pre-processing and define the metrics with which we analyzed its tags and ratings.

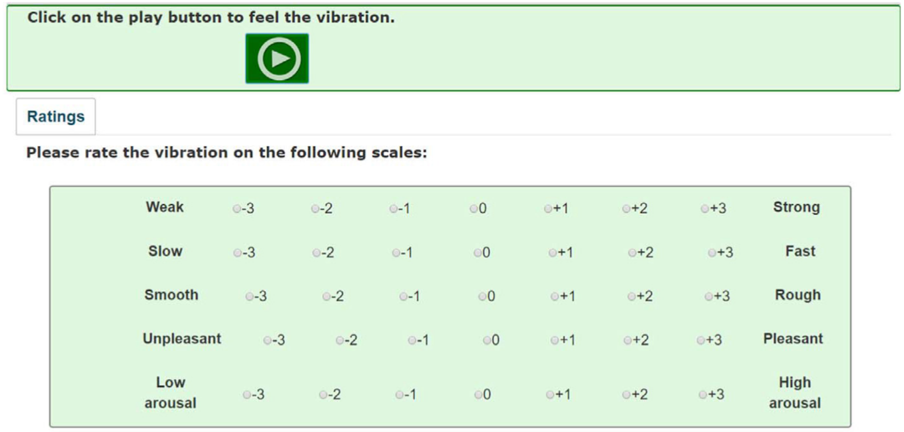
4.1. Stage 1: annotation by haptic experts

We required expert annotators who had experience with a wide range of haptic and/or vibrotactile sensations, were familiar with our vibrotactile library and facets, and could commit to annotate all or a large subset of the vibrotactile library within a short time span of a few days. Within-subject annotation of the entire vibration set would produce consistency and breadth in our initial annotation dataset; however it did impose a substantial cognitive load on the expert annotators, and thus we utilized experts with some commitment to the research and group. Given the nature of the task, we did not feel this closeness to the research could bias the results, but leveraged it for motivation.

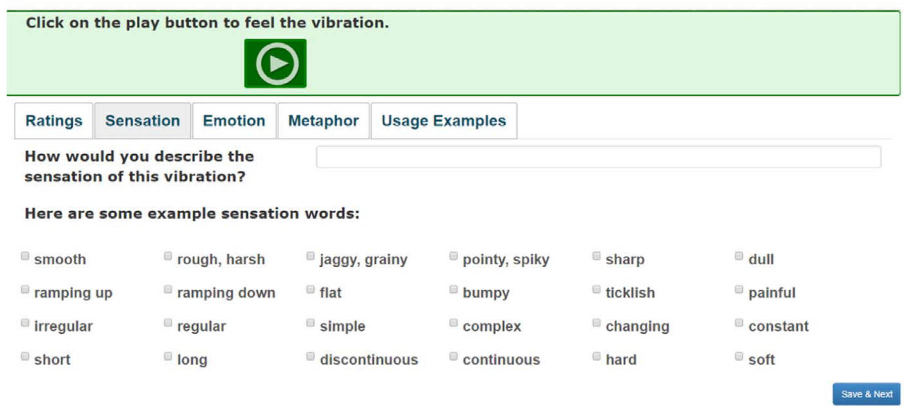
Expert backgrounds: Three haptic researchers including the first author provided expert annotations. The first author, a vibrotactile researcher who developed the vibration library and annotation instrument, rated and tagged all the vibrations while the second and third experts each annotated half of the vibrations (randomly assigned to them). The second annotator is a haptic researcher at University of British Columbia with extensive experience in designing and evaluating vibrotactile sensations and haptic design tools, The last annotator is a Human-Computer-Interaction researcher who co-designed *VibViz* (Seifi et al., 2015) with the authors and had extensive exposure to all the vibrations in the library before participating in this study. The second and third annotators received a $50 honorarium for their participation.

Initial dataset: 120 vibrations from *VibViz* library were randomly divided into 10 groups with 12 vibrations in each group. These groups remained fixed for all three expert annotators.

Apparatus and procedure: The annotation interface was a web-based wizard that gradually disclosed the available ratings and tags for the vibrations on subsequent tabs. The first tab disclosed five rating scales (7-point Likert scales) for the vibrations (Fig. 3b, Table 1). The four other tabs had the list of tags for the sensation$_f$, emotion$_f$, metaphor$_f$, and usage example$_f$ facets plus a textbox for any additional tags from the experts (Fig. 3b). In each session, the experts first played a fixed set of representative vibrations for calibration purposes, then proceeded to annotate a group of 12 vibrations (presentation order randomized). During the annotation process, the experts could play a vibration



(a) The first tab shows all the five rating scales.



(b) The other four tabs show a list of potential tags in each facet and a textbox at the top for extra tags.

Fig. 3. Expert annotation interface – one can play a vibration many times and move between the tabs representing the required ratings and tags for the vibration, but they cannot go back to previous vibration(s).

several times and move between different tabs for one vibration, but they could not go back to previous vibration(s), even within that group. At the end of each group, a review page showed all the expert's ratings and tags for the vibrations which could be further edited. This procedure encouraged the experts to focus on the demanding task of annotating each vibration individually but also allowed for cross comparisons and consistency adjustments afterwards.

Annotating a group took about 45–60 minutes. Experts were given the choice of annotating their groups in a single session or spread over several sessions, but were not permitted to interrupt a single group's annotation. Expert 1, the first author, evaluated 10 groups over 5 sessions within 6 days, while Experts 2 and 3 evaluated their 5 groups over 3/8 and 4/4 sessions/days, respectively. The experts were allowed to revisit their previously annotated groups (but never did request to do so). The total time spent by each expert was approximately 8 hours for Expert 1, and 4–5 hours for Experts 2 and 3.

Pre-processing and tag consensus and consolidation: After collecting all the annotations, the first author examined all the tags for each vibration and highlighted conflicting tags (e.g., smooth tag by one expert and rough by another one, or angry vs. happy). In a follow-up session, all three experts played and felt vibrations with contradictory tags again and came to consensus on one of the conflicting tags or on removing both. Further, they could and did adjust wording (e.g., dynamic instead of changing), and combined tags under one wording (e.g., jaggy and grainy were replaced by grainy).

4.2. Stage 2: validation of the dataset by lay users

Our sole requirement for our Stage 2 participants was to have no background in haptics beyond normal everyday exposure to vibration sensations (e.g., via cellphone usage).

Participants and compensation: We recruited 44 participants (24 female, 19–60 years old, with 40 of the participants under 36 years old) through advertising on a North American university campus. All participants were university students except for three who did not declare their occupation. Participants were permitted to participate in more than one session but tag different vibrations in each session (up to a maximum of 4 sessions covering all 120 vibrations) and six participants did so. Participants were compensated $10 for a one-hour session.

Initial dataset: Our dataset was composed of the 120 vibrations with the average expert ratings and the combined and consolidated tags for each vibration, randomly divided into 12 groups of 10 vibrations. This grouping remained fixed for all the participants.

Mitigating bias and noise in the validation stage: We anticipated that the existing ratings and tags could bias participants' perception of the vibrations and/or suggest a lower need for their attention. Following literature guidelines on detecting invalid responses (Huang et al., 2012; Curran, 2016), we mitigated this by making additions to the database which would *expose* non-diligent participants, and warned participants of the possibility of inconsistencies to *encourage* diligence, while introducing minimal additional cognitive load to the annotation task.

Specifically, we included intentional errors in the dataset, duplicated some of the vibrations, and presented the existing annotations to the participants as “data from other users that can include noise”. To identify the highly-biased participants or those who did not pay close attention to the experimental task, we included two intentional errors, one in the ratings and one in the tags, in each vibration group. For the rating error, we modified the energy$_d$ rating for one of the vibrations from very high (+3 on a 7-point likert scale) to very low (−3) or vice versa. For the tags, we added an invalid tag to one of the vibrations in each group (e.g., added “long” to a vibration with the short tag) resulting in two clearly contradicting tags for the vibration. These changes were clearly in contradiction with the vibrations’ content and with other ratings and tags, and thus were easy to note after thoughtfully feeling the vibrations. Further, the number of errors in each vibration group was low; there was exactly one tag error and one rating error in each group of 11 vibrations (including 55 ratings, and 44 lists of tags). Additionally, we duplicated one of the 10 vibrations in each group (for a total of 11 vibrations) to assess the participants’ rating and tagging reliability.

Finally, as part of the experiment instructions, we told the participants that the existing ratings and tags were provided by other people and we were running this study to remove the noise in that data.

Apparatus and procedure: The validation interface was composed of two web pages, for calibration and annotation pages respectively (Fig. 4). An experiment session took about 1 hour and the participants went over 2–3 vibration groups (22–33 vibrations) depending on their annotation speed. After the initial instructions, participants went through all the calibration vibrations for that session (33 vibrations). Then, they proceeded to the annotation page where they could see all the 11 vibrations for one group (order randomized). They could change the ratings, remove tags, or add additional tags; the initial ratings and tags were visible at all times. After completing a group, the experimenter loaded the next group of vibrations and the participant went through the calibration and annotation pages for that group. At the end of the session, participants filled a short post-questionnaire for demographic information and any other relevant comments.

4.3. Pre-processing of the dataset

Prior to full analysis, we handled outliers and then averaged and incorporated our Stage 2 annotators’ input to prune tags as planned.

Outlier removal: We used participants’ performance on intentional rating and tag errors to identify outliers with high bias or low attention to the experimental task. For this purpose, we wrote a short program to extract participants’ responses on the intentional errors from the rest of the annotations and counted the number of errors that participants corrected. If a participant only modified the rating errors, we removed their tagging data and if they adjusted less than 1/3 of both the rating and tag errors, we removed all their data from the dataset. As a result, for each vibration the dataset contains data from 9 taggers and 9–13 raters (5 rating outliers, and 13 tagging outliers).

Constructing the validated dataset: To derive the validated ratings for a vibration, we averaged all the participants’ ratings for that vibration. We eliminated tags removed by more than 1/3 of the participants (≥ 4 out of 9). In this way, we removed tags that were commonly marked as irrelevant, yet did not excessively limit the dataset (to the tags approved by everyone) to allow for more interesting analysis and results.

4.4. Definition of analysis metrics

To address our research questions, we devised a set of metrics that are applicable to ratings and free-form tags and used them as the basis for our analysis. Table 2 summarizes all the metrics with mathematical formulas. Below, V_i, V_i, denote the ith vibration and its replica respectively. T_j refers to the jth tag, F_k to one of the four facets, and N_{items} to the number of items (e.g., tags, vibrations, participants). $\cap, \ominus$ denote the intersection and symmetric difference respectively of two tag sets.

5. Analysis and results

We provide our analysis procedure and results, focusing on our three research questions in turn followed by a summary of our dataset characteristics.

5.1. [Q1] Facet substructure: what are the underlying facet dimensions that dominate user reactions to vibrations?

To interpret and verify the underlying dimensions for the facets, we analyzed the data in four steps:

1. Ran a first MDS analysis on these vibration distances in each facet to determine the number of underlying dimensions for the facet;
2. Determined an initial interpretation of the dimension semantics based on frequent and contrasting tags at the ends of each dimension (Table 3);

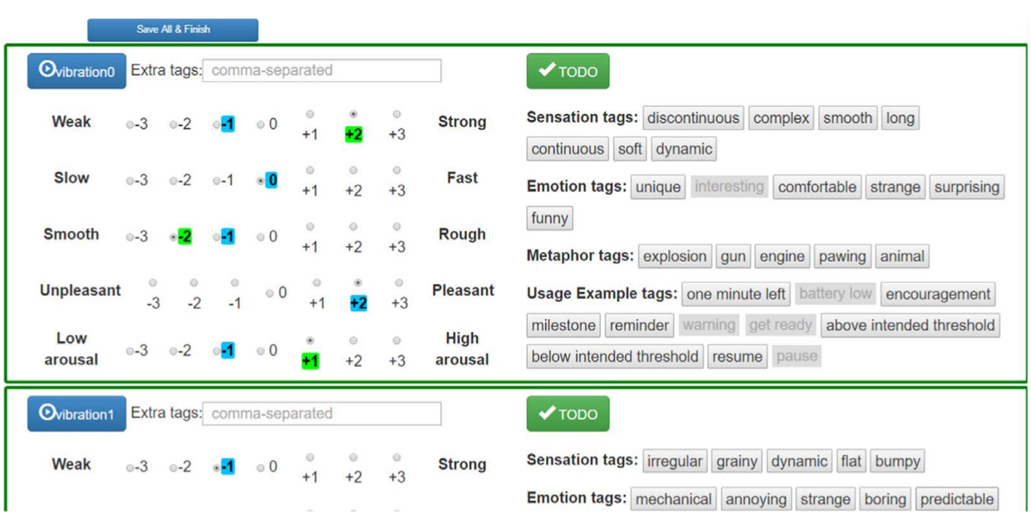


Fig. 4. The validation interface gave access to all 11 vibrations at the same time, and allowed users to remove tags and adjust ratings. Participants could see the existing (expert) ratings in blue, and their own adjusted ratings in green. They could remove a tag by clicking on it (graying it out), and re-add it by clicking it again. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Definition of our analysis metrics.

Tag removal threshold: The number of participants that must remove a tag from a vibration before we eliminate the tag in our validated dataset. For example, we use a tag removal threshold of 4, meaning that every tag that is removed by 4 or more participants from a vibration's list of tags is eliminated from the validated dataset.
Vibration distance: The extent that two vibrations are described differently according to a given metric. In our study, the metrics are our facets. We calculate the distance between two vibrations in a facet (F_k) as the number of tags (N_t) that are different between the two vibrations divided by their total number of tags in the given facet. We use this metric in our MDS analysis of the vibrations.

$$Distance(V_i, V_j, F_k) = \frac{N_{tags}[(V_i, F_k) \ominus (V_j, F_k)]}{N_{tags}(V_i, F_k) + N_{tags}(V_j, F_k)} \quad (1)$$

Tag co-occurrence and tag distance: Co-occurrence is the number of times two tags are used together to describe the vibrations in our dataset. We calculate this value for two tags by counting the number of vibrations that have both tags in our dataset and dividing it by their total frequency in our dataset.

$$Co-occurrence(T_i, T_j) = 1 - 2 \times \frac{N_{vibrations}(T_i \cap T_j)}{N_{vibrations}(T_i) + N_{vibrations}(T_j)} \quad (2)$$

Tag distance: We define distance between the two tags ("tag distance") as one minus their co-occurrence value. We use these tag distances in our MDS analysis on the tags.

$$Distance(T_i, T_j) = 1 - Co-occurrence(T_i, T_j) \quad (3)$$

Tag disagreement score: An estimate of the amount of controversy among the participants in keeping or removing a tag. We measure it based on the number participants that disagree with the majority of taggers (about removing or keeping a tag for a vibration) divided by the total number of times the tag is presented to the participants in our dataset. For example, if for all the occurrences of a tag in our dataset only one of the participant have a different opinion from the rest, the formula results in a disagreement score of 0.11. The highest disagreement is 0.44 (the lowest is 0) meaning that for all the vibrations, the tag is approved by half of the participants and removed by the other half.

$$Disagreement(T_i) = \sum_j \frac{N_{MinorityParticipants}(V_j, T_i)}{N_{vibrations}(T_i) \times N_{participants}(V_j)} \quad (4)$$

Vibration disagreement score: The amount of difference in the participants' descriptions of a vibration according to a criteria. In our study, we calculate vibration disagreement per rating and per facet. For the ratings, we use the standard deviation of the ratings by the participants. For each facet (i.e., tag set), we define our metric to be similar to the standard deviation but applicable to the tags. Specifically, for a vibration, we count the number of tags that are different between a participant's approved tags and the validated tag list for the vibration and divide it by total number of tags the experts provided for that vibration. We average the value over all taggers for that vibration.

$$Disagreement(V_i) = \sum_j \frac{N_{tags}[(V_i, P_j) \ominus (V_i, Validated)]}{N_{tags}(V_i, Experts)} \quad (5)$$

Unreliability score: *Rating* unreliability is the absolute difference in the ratings for a vibration and its duplicated version (for example, for energy ratings, the reliability is defined as $R(V_i, energy) = |energy(V_i) - energy(V\prime_i)|$). *Tag* unreliability is the percentage of removed tags that are different between a vibration and its replica. Specifically, it is the number of tags removed from a vibration or its replica (but not from both) divided by the total number of tags removed from each.

$$TagUnreliability(V_i, F_k) = \frac{N_{RemovedTags}[(V_i, F_k) \ominus (V\prime_i, F_k)]}{N_{RemovedTags}(V_i, F_k) + N_{RemovedTags}(V\prime_i, F_k)} \quad (6)$$

3. Visualized distribution of the vibrations along each MDS dimension, color-coded based on the existence (or lack) of related tags, to verify our interpretation of the dimension (Figs. 6–9);
4. Examined results of a separate MDS analysis on tag (in contrast to vibration) distances as a test of convergent and discriminant validity (Appendix D)

Together, these analyses reinforced our interpretation of the semantics of the dimensions and revealed the distribution of vibrations and tags in each facet. Below, we separately describe the analysis steps in detail, then present results for each facet.

[Step 1] Deriving dimensions from vibration distances: We calculated quantitative values for vibration distances, in each facet, based on the number of shared and different tags in the validated tag lists for each two vibrations in the library (Table 2). Then, we ran an MDS analysis on these vibration distance values for each facet. From this data, we determined the number of MDS dimensions using the eigenvalue plots as well as dimension interpretability. In Fig. 5, eigenvalue contributions are normalized to that of the first eigenvalue. Since these plots do not have an obvious "knee" (vertical gap), for each we first chose an initial set of dimensions based on the highest-contributing eigenvalues; then, considered dimension interpretability before arriving at a final number (Guest et al., 2011). We thereby found between one and four dimensions for each facet (Table 3).

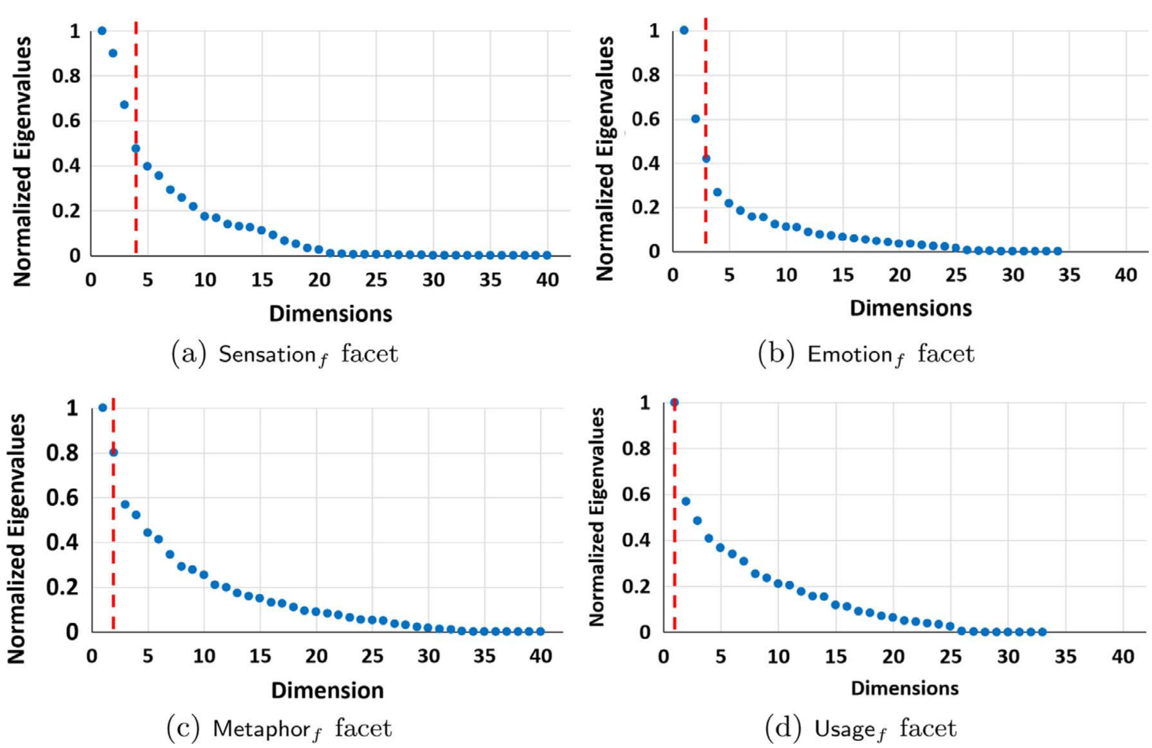
[Step 2] Determine semantic descriptors for each MDS-produced dimension: We listed the validated tags and their rate of occurrence for the 10 farthest vibrations at each end of an MDS dimension. The most frequent, yet still contrasting tags for the two ends of a dimension provided us with an initial interpretation of dimension semantics. We found one to several such *high-frequency tags* (descriptive terms) bounding each end of each dimension found in Step 1 (Table 3).

[Step 3] Verifying dimension semantics by visualizing vibration distributions: We visualized spatial distribution of vibrations along the identified MDS dimensions from Step 1 and color-coded them based on existence (red, green) or lack (gray) of high-frequency tags from Step 2 (Figs. 6–9). As explained more fully in the first caption, vertical bars encode MDS position of the vibrations along each dimension, while bar color denotes whether a vibration's validated tag list has one of that dimension's high-frequency tags. Red and green bars that are grouped at the opposite ends of the dimension with gray mostly in the middle signify that the identified tags adequately represent the semantics of the dimension; substantial mixing of colors does not.

[Step 4] Investigating tag distances: We ran a second MDS analysis on our derived *tag* distances (Table 2) and examined word positions in the resulting MDS map as a measure of convergent and discriminant validity of our interpretations (Guest et al., 2011), as follows. Convergent validity is supported when the words that have similar meanings in relation to a dimension are spatially close in the MDS solution. Discriminant validity is supported if the words with contrasting meanings are located far from each other in the MDS solution. Thus, we examined whether the contrasting tags for each dimension are far away from each other while the relevant tags for the dimensions are in the same area. Since

Table 3
Final facet dimensions (derived in Table 4) and their most frequent tags: the left column presents the number of dimensions identified from MDS analysis on the vibration distances and our interpretation of their semantics. Middle and right columns show the most frequent tags and their rates of occurrence for the 10 vibrations at the negative and positive ends of each dimension respectively.

Dimension Semantics	Negative End of Scale	Positive End of Scale
Sensation$_f$ Facet		
SensationD1:complexity$_d$	simple (8), regular (7), soft (7)	dynamic (10), irregular (9), complex (7)
SensationD2: continuity$_d$	discontinuous (10), regular (9)	continuous (10), simple (7)
SensationD3: roughness$_d$	smooth (10), soft (7), regular (7)	rough (8), short (6), discontinuous (6)
SensationD4: duration$_d$	discontinuous (7), simple (7), short (6)	grainy (8), regular (7), long (6), rough (6), ramping up (6)
Emotion$_f$ Facet		
EmotionD1: agitation$_d$	comfortable (10), calm (10), pleasant (8)	annoying (10), mechanical (9), agitating (9), urgent (9), angry (8)
EmotionD2: liveliness$_d$	predictable (10), boring (9), mechanical (9)	lively (10), unique (9), interesting (8), rhythmic (8)
EmotionD3: strangeness$_d$	rhythmic (10), lively (9), mechanical (8)	strange (10)
Metaphor$_f$ Facet		
MetaphorD1: on-off, nuanced/ongoing, repetitive$_d$	tapping (10)	engine (10)
MetaphorD2: natural/ mechanical$_d$	tapping (9), heartbeat (5)	alarm (10), game (7)
Usage$_f$ Facet		
UsageD1: urgency, attention-demand$_d$	pause (10), battery low (9), get ready (8), resume (7)	alarm (10), overtime (9), running out of time (9), above intended threshold (8)



(a) Sensation$_f$ facet

(b) Emotion$_f$ facet

(c) Metaphor$_f$ facet

(d) Usage$_f$ facet

Fig. 5. Eigenvalue plots for the four facets. In each, the horizontal axis represents number of dimensions and the vertical axis indicates a dimension's contribution to reconstructing the vibration distances. If there is a large vertical gap between the nth and (n+1)th dimensions, the first n dimensions have much larger contributions than the following ones and describe most of the variation in a facet. Thus, we use those first n dimensions in our analysis. The red dotted line highlights the number of dimensions we select for each facet. The eigenvalue contributions are normalized based on the first (largest) eigenvalue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

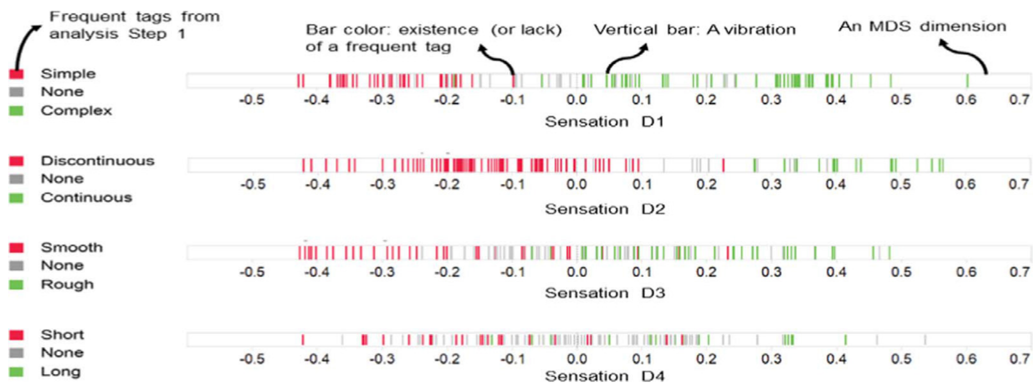


Fig. 6. Distribution of vibrations across the four MDS dimensions identified for the sensation$_f$ facet. All vibrations are shown. Position coding: thin vertical bars project each vibration's MDS-derived location onto this dimension. Color coding: bar color indicates whether the validated tag list for the vibration contains one of the frequent tags identified in Step 2 (red or green, with red indicating the left end of the dimension, and green the right end) or not (gray). For SensationD1$_d$, a red bar denotes that a vibration has a simple or a flat tag, while a green bar represents a vibration with a complex or dynamic tag and gray bars show vibrations with no related tag. SensationD2$_d$: (red: discontinuous, green: continuous), SensationD3$_d$: (red: smooth or soft, green: rough), SensationD4$_d$: (red: short, green: long). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

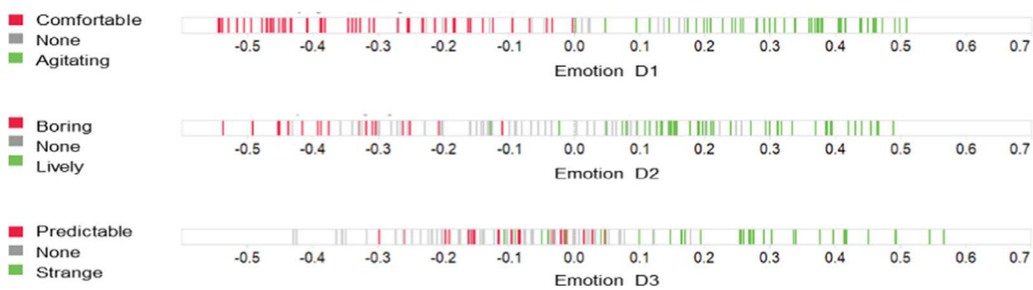


Fig. 7. Distribution of all the vibrations across the three MDS dimensions for the emotion$_f$ facet. EmotionD1$_d$: (red: calm, comfortable, or pleasant, green: urgent, annoying), EmotionD2$_d$: (red: boring, green: interesting, lively), EmotionD3$_d$: (red: predictable, familiar, green: strange, creepy, surprising). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

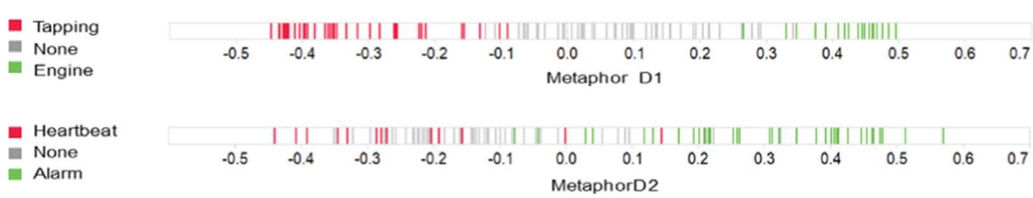


Fig. 8. Vibration distribution across the two MDS dimensions for the metaphor$_f$ facet. Tags for MetaphorD1$_d$: (red: tapping, green: engine), MetaphorD2$_d$: (red: heartbeat, green: alarm or game). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

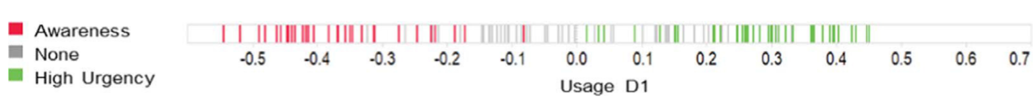


Fig. 9. Vibration distribution for the usage$_f$ facet. We color all vibrations with high urgency and attention tags (alarm, running out of time, overtime, or above intended threshold) with green marks; use red for those with awareness notifications (pause, battery low, resume, or get ready); and gray for those with none of those tags or with a mix of both types. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

this step mainly confirms the results of the above steps, we report the results in Appendix D.

In Table 4, we step through this process to interpret the dimensionality of each of our facets specifically.

5.1.1. *Our five rating scales*

To determine if our rating scales are orthogonal, we ran a Pearson correlation on the ratings for the five Likert-scale parameters across the 120 vibrations.

Results show significant medium to high correlation for all five parameters, except for pleasantness$_d$ and tempo$_d$ (low correlation, r=−0.22). Energy$_d$, arousal$_d$ and roughness$_d$ have the highest correlations (r=0.74–0.92), followed by pleasantness$_d$ and roughness$_d$ (r=−0.61), tempo$_d$ with arousal$_d$ (r=0.56), and roughness$_d$(r=0.52), and pleasantness$_d$ with arousal$_d$ (r=−0.53) (full correlation table in Appendix C).

5.2. *[RQ2] Between-facet linkages: how are attributes and dimensions linked across facets?*

We address this question by examining linkages among our identified dimensions as well as linkages among the tags between various facets.

Table 4
Facet dimension analysis.

Sensation facet

Dimensions from vibration distances: Fig. 5a's eigenvalue plot suggests that after 4 primary dimensions, additional dimensions contribute little more (< .1 apart). The identities of the most frequent tags at dimension extremes suggest that these four dimensions could be defined by their endpoints as: 1) simple/flat to complex/dynamic, 2) continuous to discontinuous, 3) smooth to rough, and 4) short to long (Table 3).

Color-coded vibration distributions: Fig. 6 shows spatial distribution of the vibrations along the above four dimensions. All four have similar ranges (−0.5 to +0.7), indicating comparable variations along the dimensions. For the first three, the associated tags explain the dimension semantics well: green and red bars are well-separated at the two ends of the dimensions and the gray bars are around the central, neutral position. For the fourth dimension, the colored bars are less well separated, suggesting that these tags can at least partially explain this variation. We include it as the last interpretable dimension for the sensation$_f$ facet. These dimensions were further confirmed by our MDS analysis on tag distances (Appendix D).

Final dimensions (also in Table 3): 1) simple—complex$_d$, 2) discontinuous—continuous$_d$, 3) smooth—rough$_d$, and 4) short—long$_d$. The overlap in the frequent tags for different dimensions (Table 3) and their spatial configuration (Fig. D.13) suggest the above dimension properties are not completely orthogonal.

Emotion facet

Dimensions from vibration distances: Fig. 5b's eigenvalues suggest 3–4 underlying dimensions; we opt for three due to higher interpretability. The most frequent tags in Table 3 suggest 1) comfortable and calm vs. annoying and urgent, 2) boring and predictable vs. lively and interesting, 3) strange and surprising vs. rhythmic and mechanical.

Color-coded vibration distributions: Fig. 7 shows the distribution of the vibrations along each emotion$_f$ dimension. For the first and second, color distribution follows our interpretation. For the last, green bars are mostly grouped at the right (strange and surprising) but red and gray bars are randomly dispersed on the left, suggesting the need for a better description for this end of the dimension.

*Final dimensions:*1) comfortable—urgent, agitating$_d$, 2) boring—lively, interesting$_d$, 3) creepy, strange—rhythmic, predictable$_d$.

Metaphor facet

Dimensions from vibration distances: We removed 13 of 45 metaphor$_f$ tags that were applied with low frequency (to < 2 vibrations) to avoid unrepresentative distortions in the MDS result. Metaphor$_f$'s eigenvalue plot then has a large number of dimensions with similar contributions; however, the first two slightly more than others. Tag frequencies suggest that these two are differentiated in 1) tapping vs. engine, and 2) tapping and heartbeat vs. game or alarm. Further tag analysis (Appendix D) indicated that along dimension 1, tags are divided into ongoing/repetitive or pulse-like/nuanced. For dimension 2, tags tend to be natural and calm; or mechanical, synthetic and annoying (see Appendix D for the spatial configuration of tags).

Color-coded vibration distribution: Tag distributions for both dimensions show a separation of green and red bars at the two ends of the dimensions with gray bars lying mostly in the middle (Fig. 8).

Final dimensions: 1) on-off, nuanced—ongoing and repetitive$_d$ metaphors, and 2) natural, calm (mostly pulsing)—mechanical and annoying$_d$ metaphors.

Usage facet

Dimensions from vibration distances: Eigenvalues suggest that the first dimension has a dominant contribution (Fig. 5d). According to the most frequent tags, this dimension represents urgency and attention-demand of notifications. On one end, usage$_f$ tags suggest time urgency while on the other, notifications require little attention and are mostly for users' awareness (Table 3).

Color-coded vibration distribution: In Fig. 9, red, gray, and green bars are fairly well separated and gradually change from the left to the right of the dimension, supporting our one-dimension interpretation for the usage$_f$ facet.

Final dimension: 1) Low-demand awareness—urgent and attention-demanding$_d$ notifications.

5.2.1. Dimension level: are there linkages or correlations among the identified dimensions of various facets? What factors can describe these correlations?

To address this question, we use factor analysis. Here, we include both ratings and facet dimensions in our analysis to further link our derived facet structures to one another as well as to the rating metrics frequently used in the literature. Thus, our variables are the five rating scales and the 10 dimensions identified for all the facets (a total of 15 variables). We use the values of the 120 vibrations on those 15 variables as our samples. This results in a ratio of 8:1 for our analysis (8 samples

Table 5
Factor analysis outcome. The left-most column shows the initial rating scales † and new facet dimensions after MDS analysis. The next four columns present the factors upon which we have found some degree of cross-facet correlation, in terms of facet ratings and dimensions. For each factor column, boldfaced numbers highlight facet variables with the highest contributions to that factor (> .45); empty cells indicate very low contributions (< .3). Facet properties that have high values on the same factor column (e.g., energy, arousal, UsageD1$_d$ in the Urgency$_f$ factor) are correlated: the columns/factors are a point of linkages between the respective facets.

Revised facet properties	Urgency (Factor 1)	Liveliness (Factor 2)	Roughness (Factor 3)	Novelty (Factor 4)
1. Sensation$_f$:				
energy$_d$†	**0.89**			
tempo/speed$_d$†	0.43	**0.45**	0.34	
roughness$_d$†	**0.75**		**0.48**	
SensationD1 - Complexity$_d$	**0.45**			**0.55**
SensationD2 - Continuity$_d$		-0.38		0.31
SensationD3 - Roughness$_d$			**0.89**	
SensationD4 - Duration$_d$	0.36	**-0.48**		
2. Emotion$_f$:				
pleasantness$_d$†	**-0.64**	0.33	-0.34	-0.31
arousal$_d$†	**0.95**			
EmotionD1 - Agitation$_d$	**0.82**			
EmotionD2 - Liveliness$_d$		**0.89**		
EmotionD3 - Strangeness$_d$				**0.60**
3. Metaphor$_f$:				
MetaphorD1 - On/off vs. ongoing$_d$		-0.32		0.44
MetaphorD2 - Natural vs. Mechanical$_d$	**0.45**			
4. Usage$_f$:				
UsageD1-Attention-demand$_d$	**0.80**			

Sensation tags

Emotion tags

	bumpy	complex	cont.	discont.	dynamic	firm	flat	grainy	irregular	long	pointy	rampdwn	rampup	regular	rough	short	simple	smooth	soft	spiky	springy	ticklish	wavy
agitating	0.16	0.36	0.16	0.46	0.39	0.21	0.04	0.29	0.36	0.2	0	0	0.23	0.3	0.67	0.3	0.08	0.06	0	0.32	0	0.24	0
angry	0.11	0.1	0.14	0.26	0.25	0.21	0	0.16	0.17	0.15	0	0.07	0.19	0.19	0.5	0.19	0.04	0	0	0.26	0	0.09	0
annoying	0.16	0.24	0.34	0.35	0.42	0.22	0.09	0.39	0.25	0.21	0	0.09	0.37	0.24	0.66	0.31	0.11	0.03	0.03	0.22	0	0.15	0
boring	0.14	0	0.37	0.13	0.14	0.14	0.47	0.36	0.04	0.24	0	0.2	0.09	0.24	0.11	0.19	0.39	0.28	0.27	0.05	0	0	0.1
calm	0.23	0.06	0.22	0.39	0.15	0.08	0.43	0.2	0.09	0.32	0.05	0.27	0.13	0.48	0.05	0.25	0.54	0.62	0.58	0.1	0	0	0.05
comfortable	0.5	0.21	0.26	0.52	0.37	0.03	0.22	0.26	0.27	0.21	0.04	0.19	0.16	0.4	0.09	0.26	0.4	0.55	0.56	0.08	0.04	0	0.07
creepy	0	0.13	0.13	0.02	0.07	0	0.12	0.11	0.06	0.07	0	0	0.13	0	0.05	0.07	0.05	0	0	0	0	0.2	0
energetic	0.44	0.1	0	0.29	0.11	0.08	0	0.13	0.04	0.1	0	0	0.1	0.31	0.04	0.2	0.26	0	0.1	0.28	0.13	0	0
familiar	0.29	0	0.19	0.31	0.05	0.1	0.2	0.17	0	0.31	0	0.2	0.08	0.45	0.06	0.26	0.6	0.37	0.48	0.12	0	0	0
fear	0.04	0.16	0.11	0.18	0.21	0.17	0.08	0.23	0.29	0.17	0	0	0.16	0.12	0.27	0.22	0.04	0	0	0.12	0	0.35	0
funny	0.08	0.3	0.06	0.12	0.19	0	0	0.15	0.21	0	0	0	0.06	0	0.13	0.12	0.04	0.05	0.06	0.14	0.22	0.15	0.2
happy	0.41	0.18	0.05	0.31	0.21	0	0.06	0.2	0.12	0.14	0	0	0.05	0.34	0.1	0.09	0.21	0.2	0.22	0.15	0.1	0	0.1
interesting	0.48	0.46	0.14	0.45	0.46	0.05	0	0.16	0.33	0.15	0	0	0.21	0.23	0.2	0.21	0.23	0.22	0.14	0.23	0.06	0.06	0
lively	0.54	0.33	0.03	0.58	0.39	0.12	0	0.16	0.34	0.18	0	0.04	0.18	0.41	0.2	0.24	0.3	0.22	0.18	0.32	0.05	0.04	0.05
mechanical	0.24	0.32	0.36	0.58	0.51	0.2	0.1	0.44	0.4	0.26	0.03	0.17	0.4	0.41	0.61	0.28	0.28	0.16	0.08	0.24	0	0.08	0
natural	0.3	0.09	0.09	0.27	0.13	0.06	0.13	0.12	0.08	0.14	0	0.13	0	0.34	0	0.18	0.34	0.46	0.43	0.1	0	0	0.09
pleasant	0.51	0.18	0.11	0.38	0.18	0	0.14	0.25	0.1	0.15	0	0.05	0.07	0.41	0.03	0.25	0.49	0.44	0.56	0.15	0.06	0	0.12
predictable	0.22	0.03	0.41	0.23	0.22	0.09	0.36	0.37	0.06	0.32	0	0.13	0.24	0.44	0.19	0.14	0.5	0.4	0.37	0.07	0	0	0.06
rhythmic	0.55	0.34	0.03	0.54	0.29	0.2	0.04	0.08	0.34	0.25	0.05	0.08	0.03	0.43	0.23	0.12	0.28	0.22	0.21	0.26	0.05	0.04	0
sad	0.04	0.06	0.24	0.08	0.09	0.11	0.5	0.05	0	0.31	0	0.4	0.18	0.15	0	0.06	0.26	0.3	0.24	0	0	0	0
strange	0.22	0.39	0.28	0.33	0.43	0.05	0.14	0.28	0.45	0.14	0	0.23	0.28	0.11	0.23	0.25	0.2	0.16	0.1	0.15	0	0.05	0
surprising	0.13	0.48	0.26	0.29	0.47	0.19	0	0.15	0.43	0.04	0	0.18	0.35	0.03	0.33	0.17	0.03	0.11	0.04	0.14	0.09	0.15	0
uncomfortable	0	0.29	0.21	0.27	0.28	0.35	0.06	0.25	0.15	0.21	0	0.17	0.42	0.13	0.45	0.25	0.1	0.04	0	0.14	0	0.21	0
unique	0.28	0.68	0.18	0.41	0.55	0.05	0.05	0.16	0.52	0.18	0	0.09	0.14	0.14	0.23	0.14	0.11	0.22	0.11	0.27	0.06	0.11	0
urgent	0.19	0.37	0.29	0.49	0.46	0.25	0.04	0.34	0.35	0.23	0	0.11	0.34	0.31	0.71	0.26	0.17	0.08	0	0.27	0	0.16	0

Fig. 10. Co-occurrence of the sensation$_f$ tags with the emotion$_f$ tags in our vibration library. For each emotion$_f$ tag (rows), we see the most (and least) associated sensation$_f$ tags (encoded as darker and lighter cells respectively). For example, highlighted with red boxes, to design a surprising vibration, one should make an irregular, dynamic, ramping up, and rough sensation (design scenario in Fig. 1a). Similarly, looking down on each column, one can see how a particular sensation$_f$ tag is perceived emotionally. Bumpy vibrations mostly invoke positive emotional response such as comfortable, energetic, happy, lively, etc. (evaluation scenario in Fig. 1b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

per variable), satisfying the minimum suggested ratio in the literature (5:1) (Yong and Pearce, 2013).

According to our results, four perceptual factors can describe correlations among the dimensions in various facets (the four right-most columns on Table 5). Table 5 shows the vibration properties (ratings and facet dimensions) with loadings >0.3 for each factor and highlights the ones with higher loadings (≥4.5) in **boldface**.

Factor 1 (Urgency$_{fact}$): UsageD1$_d$ and emotionD1$_d$ are highly connected to the same underlying factor as energy$_d$, arousal$_d$, roughness$_d$, and pleasantness$_d$. SensationD1 - complexity$_d$ and metaphorD2 - natural/mechanical$_d$ are also connected to this factor but with lower loadings.

Factor 2 (Liveliness/interestingness$_{fact}$): EmotionD2 – boring/lively$_d$ is connected with sensationD4 – duration$_d$, and tempo$_d$ on the second factor. SensationD2 – continuity$_d$ is also partially loaded onto this factor.

Factor 3 (Roughness$_{fact}$): This factor presents the link between sensationD3 – roughness$_d$ with roughness$_d$ and pleasantness$_d$ ratings.

Factor 4 (Novelty$_{fact}$): SensationD1 – complexity$_d$ and emotionD3 – strangeness$_d$ are connected on the fourth factor. MetaphorD1$_d$ also partially loads onto this factor.

5.2.2. Tag level: how do tags in the different facets correlate?

We used our tag co-occurrence metric (Section 2) as a measure of correlation between tags in various facets. We report co-occurrence of the sensation$_f$ facet's tags with emotion$_f$, metaphor$_f$, and usage$_f$ tags, since sensation$_f$ tags more directly relate to engineering parameters (Fig. 10) but are also hardware independent. Fig. 10 presents links among the emotion$_f$ and sensation$_f$ tags (see Appendix F for the tag co-occurrence tables of the metaphor$_f$ and usage$_f$ facets).

5.3. [RQ3] Individual differences: to what extent do people coincide or differ in their assessment of vibration attributes?

We examined variation in the participants' ratings and tags as a measure of individual differences in their perceptions and opinions. Here, we report these individual differences on various levels including the extent of variation (disagreement) across the facets, ratings, and tags as well as the amount of disagreement per vibration.

5.3.1. Per facet

We measured overall individual differences in the facets based on percentage of facet tags that were approved by everyone (100% of the annotators), as well as percentage of tags that caused a split between the participants (defined as when half of participants removed a tag and the other half kept it as an appropriate tag for a vibration). Sensation$_f$ had the lowest I.D.$_f$, with the highest number of tags kept by everyone (21% compared to 7–12% for the other facets), and the lowest number of tags that caused a split (18% compared to 32–37%). Usage$_f$ elicited slightly more individuated responses than emotion$_f$ and metaphor$_f$, with 7% tags approved by everyone and 37% tags resulting in a split in the participants' opinions.

5.3.2. Per rating

For each of the five rating scales, we used standard deviation of the values provided by all the annotators for a vibration as a measure of I.D. in that rating. Averaged across all vibrations and on a 7-point scale, these

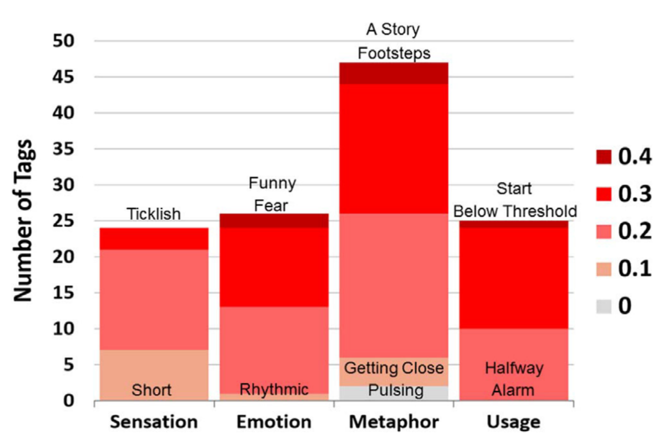


Fig. 11. A stacked bar chart showing tag disagreement scores in each facet. The height of each bar indicates total number of tags in a facet. More saturated parts of the bar indicate tags with higher disagreement scores. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

	Energy	Tempo	Roughness	Valence	Arousal	Sensation	Emotion	Metaphor	Usage
v-09-10-3-56	0.44	0.62	0.59	0.91	0.79	0.02	0.39	0.33	0.29
v-09-10-4-25	0.35	0.59	1.26	0.96	0.81	0.3	0.43	0.07	0.28
v-09-10-6-46	0.4	0.4	0.67	0.91	0.84	0.27	0.2	0.44	0.37
v-09-12-1-0	0.46	1.39	0.86	1.07	0.83	0.27	0.41	0.24	0.25
v-10-28-7-35	0.99	1.49	1.16	1.19	0.53	0.16	0.31	0.41	0.4
v-09-09-8-11	1.08	0.54	0.59	0.63	0.5	0.15	0.22	0.22	0.27
v-09-09-8-20	0.92	0.65	1.1	0.91	0.76	0.22	0.17	0.19	0.3
v-09-09-8-20-cpy	0.85	0.36	0.85	0.76	0.69	0.22	0.14	0.19	0.21
v-09-09-8-24	0.89	1.1	1.11	1.38	0.62	0.2	0.36	0	0.33
v-09-10-11-55	0.47	0.96	1.1	1.01	0.83	0.09	0.17	0.22	0.23
v-09-10-12-11	0.57	0.53	0.28	0.99	0.59	0.15	0.09	0.11	0.06
v-09-10-12-13	0.28	0.58	0.65	1.18	0.38	0.09	0.26	0.14	0.26
v-09-10-12-16	0.4	1.11	0.79	1.33	0.44	0.17	0.27	0.44	0.24
v-09-10-12-2	0.79	0.36	0.52	0.86	0.28	0.16	0.2	0.22	0.22
v-09-10-12-6	0.77	0.62	0.46	0.93	1.01	0.03	0.17	0.16	0.27
v-09-10-12-9	0.53	0.5	0.66	0.88	0.18	0.21	0.26	0.27	0.28
v-09-10-12-9-cpy	0.68	0.33	0.73	1.11	0.66	0.24	0.19	0.38	0.39
v-09-10-3-52	0.14	0.26	0.57	0.83	0.64	0.2	0.28	0.24	0.22

Fig. 12. Disagreement scores for the ratings and facets for a subset of the vibrations, calculated based on Table 2. Disagreement scores are within [1–7] (ratings), and [0–1] (facets). A vibration can have a low disagreement score on one rating or tag set but a high disagreement score on another.

are 1.0, 0.8, 0.7, 0.7, 0.7 for pleasantness$_d$, roughness$_d$, energy$_d$, tempo$_d$, and arousal$_d$ respectively.

5.3.3. Per tag

Stage 2 participants approved or removed some tags in consistent ways (e.g., short, irregular, agitating) whereas the participants showed differing opinions about the appropriateness of some others (e.g., ticklish, fear, start). *Tag disagreement score* represents the amount of controversy among the participants in keeping or removing a tag (Section 4.4). The highest possible score is 0.5, denoting a full split in participant opinions.

Fig. 11 shows a bar chart of the number of tags in each facet, color-coded based on their disagreement score (higher color saturation denote higher disagreement scores). The figure also lists examples of tags with low and high disagreement scores: e.g., in sensation$_f$, short and smooth transition tags had the lowest disagreement while ticklish had the highest. Overall, usage$_f$ tags had higher disagreement compared to the other facets, with no tag showing very low (< .2) disagreement.

5.3.4. Per vibration

We computed disagreement among the ratings and tags assigned to each vibration (vibration disagreement score is defined in Section 4.4). Fig. 12 presents a heatmap of a subset of vibrations and their disagreement scores for the ratings and tags (see disagreement values for all the vibrations in Appendix E.17). Interestingly, the vibrations were not always consistently disagreed or agreed upon. For example, vibration “v-09-10-3-56” had low disagreement on sensation$_f$ tags but higher disagreement on emotion$_f$, metaphor$_f$, and usage$_f$ tags. The vibrations also differed in the facet(s) that had the *lowest* controversy for them: “v-09-10-6-46” was mostly agreed upon in the emotion$_f$ facet but had high disagreement in the metaphor$_f$ facet. This pattern was reversed for another vibration (e.g., “v-09-10-4-25”).

Table 6
Summary of our annotation dataset after the two stages of expert annotation and lay user validation (i.e., pruning). The left column indicates: the average difference in values provided on the five rating scales originally used to define the facets (top section); overlap in the tag sets for each of the facets (middle section); and the overall tag count for these facets (bottom section). Values in the experts and lay user columns in Table 6 cannot be directly compared due to differences in the tasks in these collection stages: experts applied annotations (each vibration was annotated by two of three experts), while lay users were asked to confirm them, and largely removed rather than adding tags.

	Experts Average difference among experts	Lay Users Average deviation from experts
Rating difference	(Range, 7-point scale)	(Range, 7-point scale)
Energy$_d$	1.15	0.45
Tempo$_d$	1.26	0.54
Roughness$_d$	1.6	0.64
Pleasantness$_d$	1.64	0.84
Arousal$_d$	1.5	0.5
Tag overlap	Tags applied by both experts	Tags approved by $\geq$ 4 lay users
Sensation$_f$	25%	86%
Emotion$_f$	17%	72%
Metaphor$_f$	14.5%	76%
Usage$_f$	12.5%	69%
Dataset tag count	Following expert annotation	Following lay-user validation
Sensation$_f$	744	635
Emotion$_f$	988	716
Metaphor$_f$	584	442
Usage$_f$	1234	857

5.4. Methodology: how does staged data collection impact annotation quality?

The goal of our two-stage data collection was to reduce noise from outliers and improve dataset convergence and reliability by facilitating the annotation task for the lay users, but at the cost of an additional round of data collection. Below, we summarize how well this new method achieves these goals by examining dataset characteristics after the two rounds of annotations and reliability of the final dataset.

Expert and Lay User Annotations: Table 6 summarizes characteristics of our dataset after expert and lay user annotation stages.

Reliability of the final annotation set: To assess reliability, we measured absolute rating difference and percentage of tag difference between a vibration and its replica (Section 4.4) for each individual participant as well as for the final aggregated dataset. On average, the ratings were ~ .7 apart (on a 7-point scale) for individual participants but this difference was reduced to ~ .2 for the final aggregated dataset. Further, ~ 33% of the tags removed by an individual were different between a vibration and its replica which was further reduced to ~ 7% difference on the final aggregated set.

6. Discussion

We start by looking at how these results apply to the three design, evaluation, and personalization scenarios we proposed in the introduction (Fig. 1): have we indeed found evidence for *perceptually continuous dimensions within individual facets, along which users would presumably find it logical to "move" individual haptic elements as an act of design?* Do we have a mapping among the facets that enables translation of design requirements, or evaluation of aesthetic properties of haptic elements?

We then compare our facet dimensions with the perceptual vibrotactile properties in the literature and draw insights into findings on individual differences and annotation reliability. We finish by reviewing the validity and effectiveness of our methodological choices.

6.1. Within-facet perceptual continuity: scenarios

Scenario 1 – design guidelines and manipulations (Fig. 1a): In making haptic sensations, designers commonly have a set of requirements in the usage$_f$, metaphor$_f$, or emotion$_f$ facets (e.g., surprise or racing car engine) and require guidelines prescribing important sensation$_f$ or engineering parameters for meeting those requirements. The linkages between the facets can provide such guidelines: the designer can look along the rows of Fig. 10 and find the highly correlated sensation$_f$ tags. For example, using Fig. 10, the task of designing a surprise vibration is broken into designing a sensation that is irregular, complex, ramping up, and rough (sensation$_f$ tags with high co-occurrence with surprise).

On the dimensional level, between-facet linkages provide a more continuous mapping for design. For example, a designer might want to create a palette of sensations that vary in liveliness. Using the correlation among the boring—lively$_d$ dimension and the dimensions from the sensation$_f$ facet, the designer can vary continuity$_d$ and tempo$_d$ of the vibrotactile rhythm in sketching alternative palettes for further investigation.

Here, we link usage$_f$, metaphor$_f$, or emotion$_f$ facets to the sensation$_f$ attributes, rather than engineering parameters of vibrations, to provide concrete yet hardware-independent design guidelines. The specific engineering parameters for the facet descriptions could vary depending on the hardware and form factor of a target device and are not generalizable. In contrast, sensory properties of vibrations (e.g., roughness, continuity, and duration) are hardware-independent descriptions, well-studied in the literature (e.g., roughness), or straightforward to implement (e.g., continuity). For example, the designer can add discontinuity by including silence or pause in a vibration while ensuring that the duration of silence is perceptible to people (Ternes, 2007).

Scenario 2 – evaluation (Fig. 1b): Alternatively, for cases where a designer has a set of vibrations and is interested to know their emotional connotations, proper metaphors or usage examples, he/she can look them up along the columns of Fig. 10. For example, a bumpy sensation usually has positive emotional connotations such as happy, interesting, lively and rhythmic, while ramping up sensations are usually annoying, mechanical, and uncomfortable.

Scenario 3 – personalization (Fig. 1c): Facet dimensions and their linkages provide the theoretical grounding for designers to *build tuning and stylization tools* for end-users who may wish to personalize their vibration notifications. First, the dimensions we found in this work are good candidates for being the basis of tuning sliders, as they capture the dominant spectrums along which a vibration can vary in a facet. For example, one can imagine a tuning slider that moves a vibration along the emotion dimension of boring—lively$_d$. Then, even more practically, the linkages identified in our results from a dimension in the emotion$_f$, metaphor$_f$, or usage$_f$ facets to the sensation$_f$ dimensions inform us about the mechanics of building these sliders. For example, the boring—lively$_d$ dimension is correlated with the signal's tempo, duration$_d$ (sensationD4) and continuity$_d$ (sensationD2). Thus, a designer can use these three sensation$_f$ attributes in developing an automated algorithm for a liveliness slider, which is ultimately controlled by end-users to modify a vibration's liveliness for their personal taste. Our future work focuses on using our results to build a set of tuning sliders for vibrations.

6.2. Facet dimensions and linkages

Here, we discuss the unique insights and challenges for the facet dimensions and present implications for future research and design when applicable.

Sensation $_f$ provides designers with a practical translation platform between the facet space and engineering parameters like frequency and waveform. Sensation$_f$ dimensions reflect important perceptual and engineering parameters identified in past studies. Specifically, *rhythm* and *envelope*, two parameters found to be influential and manipulable in expressive vibrotactile design (MacLean, 2008; Ternes, 2007), are directly linked to continuity$_d$ and complexity$_d$ (sensationD2, D1 respectively). Roughness$_d$ and duration$_d$ are also known to impact users' perception (Hoggan and Brewster, 2007a, 2007b; MacLean, 2008). Thus, translating the emotion$_f$, metaphor$_f$, and usage$_f$ dimensions and tags to the sensation$_f$ facet offers a practical and hardware-independent means for design.

Emotional perceptions of vibrations do not follow theoretical dimensions of pleasantness $_d$ and arousal $_d$. Correlation of the pleasantness$_d$ and arousal$_d$ ratings (Section 5.1.1) as well as our MDS results on the emotion$_f$ tags suggest that these two dimensions are not orthogonal for our vibrotactile collection. As a result, not all four quadrants of the pleasantness (valence)-arousal grid are covered by the vibrotactile sensations in our library. Specifically, none were marked as either very pleasant and alarming (positive valence-positive arousal), or very calm but unpleasant (negative valence-negative arousal).

While it is possible that such examples exist but our library does not contain them, we note that two recent studies found a similar correlation and also the same gap for different vibrotactile actuators and vibration sets (Seifi and MacLean, 2013; Yoo et al., 2015). Yoo et al. examined several sets of vibrations (24–36 items each) on a voice coil actuator (Haptuator – TactileLabs, 2016) and none covered the negative valence-negative arousal or very high valence-high arousal quadrants (Yoo et al., 2015). Our own previous study reports a similar correlation for a small subset of 14 vibrations on an Electro-Active Polymer (EAP) actuator (Seifi and MacLean, 2013).

We propose that for vibrations, the theoretical dimensions of pleasantness$_d$ and arousal$_d$ in the literature are not good representatives for the 2-D affect grid. There, sad and boring have negative valence and negative arousal while vibrations with sad and boring tags do not fall in that area; they are not necessarily unpleasant and quiet and this difference is reflected in our dataset. Instead, our MDS analysis on the emotion$_f$ tags suggest that people perceive and rate vibrations according to three other dimensions: 1) agitation$_d$, 2) liveliness$_d$, and 3) strangeness$_d$.

This result impacts future research and design in at least three ways. First, further studies are needed to *confirm or reject this pattern*, and compare emotion$_f$ dimensions for vibrations with other haptic stimuli (such as natural textures, force feedback and variable friction) and other modalities such as vision and audition. Second, the three dimensions provide *new directions for vibration design*. Agitation$_d$, liveliness$_d$, and strangeness$_d$ explain large variations in emotion$_f$, have low correlation, and provide a more accessible design space for current vibrotactile technology. They may be promising targets for affective design. Finally, once further validated, these dimensions offer good candidates for devising a standard *evaluation instrument* for vibrations.

Metaphor $_f$ dimensions are the most difficult to interpret. Our results suggest two dimensions for metaphor$_f$ tags that vary on continuity, novelty, and urgency. However, the spatial configuration of tags in Fig. D.15 does not completely follow this definition (see the report of outlier tags in Appendix D). Also, these two dimensions are partially linked to the other facets in our factor analysis. One reason could be that our metaphor$_f$ tag set is larger but also sparser: there are fewer common metaphor$_f$ tags among the vibrations (Table 3) compared to sensation$_f$, emotion$_f$, and usage$_f$ tags. While this trend can reflect an inherent characteristic of metaphors for describing vibrations, future studies are needed to validate and expand on the above dimensions and further develop the metaphor$_f$ vocabulary for vibrotactile effects by studying a broader list of initial metaphor tags collected from diverse participants.

Users' interpretation of vibration meaning in usage contexts is mainly dictated by their energy (or urgency). According to our MDS results, vibration energy$_d$ or urgency$_d$ is the most important dimension for usage$_f$ tags. While energy is an important design parameter, we are not aware of previous work that empirically connects a vibration's energy to its application. Our vibration library is designed to include a wide range of sensations but our tag list for usage$_f$ is developed for a specific context: applications where time tracking is an important component (e.g., giving presentations and exercising). We anticipate this finding to extend to other application contexts but future studies are needed to confirm or reject the importance of energy for other types of applications.

Emotional connotations of vibrations play an important role in users' perception of vibrations, regardless of facet. The three dimensions found for emotion$_f$ have substantially high loadings on three of the four factors in Table 5: urgency$_{fact}$, liveliness$_{fact}$ and novelty$_{fact}$. This suggests that the underlying constructs, describing the variations and linkages between the facets, are mainly emotional. In the absence of other strong criteria, the emotion$_f$ facet can serve as the best default for end-user tools and interfaces.

6.3. Individuals' annotation reliability and variation

Reliability of individuals' tagging is surprisingly low. In our Stage 2 study component, we placed a duplicate vibration in each vibration set – i.e., two out of the 11 were identical (Section 4.2). However, about 33% of individuals' removed tags differed for these duplicates (Section 5.4). This number is unexpectedly high: participants had access to all the vibrations and their tags via the experiment interface. Although the variation may be partially due to varying commitment and focus, it also suggests that people's memory of vibrations quickly fades. In contrast to auditory and visual icons, sensations in this unfamiliar modality are not always immediately memorable, and users commonly play a vibration several times to form an opinion about it or to compare it with another vibrotactile sensation. This negatively impacts reliability, but in some cases can simplify study design when one stimulus is presented in multiple experimental conditions.

Data on individual differences in ratings and tags inform haptic evaluation. Disagreement scores for the tags and ratings suggest that a notable portion of annotation variation is due to differences among users' definitions of the language terms and its manifestation in a tactile signal. This is evidenced by lower individual difference values for sensation$_f$ tags and the five rating scales. To mitigate this in the long run, we need to devise and consistently use a set of standard rating scales; the facet dimensions are promising candidates for such an endeavor. In the meantime, our tag disagreement scores can inform haptic researchers in selecting less controversial tags or estimating the number of participants required for their evaluation.

6.4. Review of our methodology

We contribute a data collection and analysis methodology, based on existing practices in the music annotation domain, that allows for comprehensive evaluation of a large vibration collection. Here, we discuss the validity and effectiveness of our methodological choices according to our results to support future uses and adaptations of our approach.

6.4.1. Method validity

Bias in validation stage: Seeing existing annotations did not override participant perceptions. Participants made large adjustments (~ 4.3 on a 7-point scale) to the intentional energy rating errors applied in the validation stage to identify outliers – (Section 4.2). Also, a notable percentage of the tags (~14–31%) are removed by 4 or more (out of 9) participants, demonstrating some degree of inter-participant consistency as well as willingness to respond with initiative. We also guarded against bias by describing the existing annotations to the participants as "noisy data from other users;" and by eliminating the participants with few annotation adjustments as outliers, on presumption that this indicated low engagement with the task. Finally, our validation task resembles practical scenarios where users start from a proposed set of notifications and their intended perception and usage (e.g., list of alarm tones on a phone, game sounds, etc.) and adopt or reject notifications depending on their perceptual match. Thus, although we expect some degree of conformity among the participants to the existing tags and ratings which were their (nonzero) starting point, it appears this did not override their choices and our validated dataset reflects their accepted annotations among the proposed ones.

Annotation instrument: Quality of our ratings and tag lists are reflected in our results. We included five Likert-type ratings to link our results to the most prominent vibration properties in the literature. Alternatively, these properties could be represented as tags in our instrument. While the latter would result in a uniform instrument with only tags, the former enabled us to capture users' perception of these properties with more precision and informed design of the *VibViz* interface (Seifi et al., 2015). Further, since factor analysis is robust to metrics with different scales (Yong and Pearce, 2013), we could include both MDS dimensions and ratings in one analysis, complementing our picture of users' perception of vibrations.

While developing the tag sets, our goal was to include as many relevant tags as possible, yet avoid redundant tags. For sensation$_f$ and emotion$_f$, our tag lists were built on existing adjective lists in the literature, were inclusive and were independent of the context. Thus, for these facets we could identify several dimensions with stronger linkages in the factor analysis. In contrast, the metaphor$_f$ and usage$_f$ tag lists were use-case dependent and could not be inclusive in nature. Further, it was more difficult to identify tag redundancy and conflicts for them. Thus, they resulted in fewer dominant dimensions which were harder to interpret (metaphor$_f$) and dependent on use case (usage$_f$). The attributes and dimensions for these facets can be further refined and validated over time, through follow-up studies that examine a broader range of use cases and metaphors with diverse participants.

Future work can further refine our metaphor$_f$ and usage$_f$ attributes and dimensions by studying other use cases and participant groups.

Analysis methods: We triangulate our analysis to guard against the subjectivity in our interpretations. For both MDS and factor analysis, researchers determine number and semantics of dimensions and factors. Although this interpretation is based on evidence in the data, the resulting semantics are subject to the researchers' bias and pre-conceptions. To guard against this, we use three different analyses on the tags to interpret semantics of the facet dimensions and provide data on between-facet linkages on both dimensional and tag level.

Analysis methods: Factors with low loadings must be interpreted with caution. Our factor analysis has a ratio of 8:1 for data points (120 vibration ratings and MDS positions) and variables (15 ratings and facet dimensions; Section 5.2.1). While this meets the minimum ratio proposed in the literature (5:1), higher ratios (10:1 or more) are recommended for more stable results (Yong and Pearce, 2013). With our data, the variables with low factor loadings may not be stable if more data is added, thus they must be regarded with caution. This is especially true for the two metaphor$_f$ dimensions and for continuity$_d$ (sensationD2).

6.4.2. Method effectiveness

Recruitment benefits: the staged approach increases efficiency of data collection and improves convergence. Practically speaking, we found that validating existing ratings and tags can be done more quickly than annotating a vibration. In our study, validation sessions include about three times more vibrations than our pilot and expert annotation sessions (33 vibrations compared to 12 vibrations). This means the same amount of data can be collected with fewer participants. Further, we found that the between-subject variations in the validation stage were reduced to values equal to within-subject variations (reliability) in the ratings, leading to better convergence. In Sections 5.4, and 5.3.2, all values are ≤ 1 on a 7-point Likert scale. Finally, having expert ratings on the vibrations allowed for quick detection of outliers in the data and adjusting the recruitment plan accordingly.

Value added by end-user validation: second stage is crucial for validating expert tags. On average, the lay-user-validated ratings are about 0.5 (7-point scale) different from the expert ratings, and the lay-user-validated set of tags include 14–31% fewer tags than the expert tag set. These results suggest that experts' ratings provide a fairly accurate estimate of users' ratings; while for the tags, experts' and lay participants' opinions deviate more, justifying the need for the validation stage. If further studies confirm this pattern, then this approach can provide a *discount evaluation method* for vibrotactile design similar to heuristic evaluation in user interface design (Nielsen and Molich, 1990).

7. Conclusion

Our work investigates four vibration facets, their underlying dimensions and their linkages and mappings based on ratings and tags collected for a library of 120 vibrations; Fig. 2 illustrates the emergent landscape we have exposed and described with tags, facets, dimensions and facet-linking factors. Our data and analysis confirm definite cross-facet linkages between certain facet dimensions. We describe these linkages on a discrete level between tags (descriptive words applied to specific vibrations, which themselves we have empirically located within facet dimensional space) and on a continuous level between dimensions$_d$ (wherein dimensions provide perceptual delineation of the facets). For the latter, the linkages can be described according to four factors (perceptual constructs underlying facet linkages): a vibration's urgency$_{fact}$, liveliness$_{fact}$, roughness$_{fact}$ and novelty$_{fact}$.

The linkages between the sensation$_f$ facet and the other facets (on both tag and dimension levels) offer guidelines for vibration design, evaluation, and personalization. However, we still lack a continuous mapping between most facet parameters (user's cognitive schemas) and the engineering parameters, by which these sensations are constructed. Applying machine learning techniques to the vibratory signals and their associated disposition within the facet space (such as the ratings, tags and MDS positions on the facet dimensions) is one approach towards identifying such a mapping. To this end, we have released our vibration dataset (vibration.wav files, their annotations and MDS characterization) for use by other researchers (VibViz Dataset, 2016).

Further, our lab continues to examine this mapping in the use case of developing a set of tuning sliders that can move a vibration along the semantic facet dimensions – that is, Scenario 3.

Will underlying facet dimensions and linkages apply to sensations produced with other haptic technologies? We anticipate that to a large extent they will, although specific labels and properties for the facets might vary. The literature includes evidence that people use sensation$_f$, emotion$_f$, and metaphor$_f$ descriptions for many kinds of haptic sensations, ranging from ultrahaptics effects (non-contact stimuli produced with acoustic waves (Obrist et al., 2013)) to movements of a furry touch-based social robot (Yohanan and MacLean, 2011; Yohanan et al., 2005). Confirming this requires future studies that examine the facet dimensions for other types of haptic sensations, such as force feedback, texture displays, variable friction and ultrahaptics, and comparing their findings with our results. Such an endeavor can lead to a more holistic and technology-independent model of user haptic perception.

Beyond haptics, facets can facilitate multimodal experience design. First, they support design of rich haptic experiences, which in turn enhances multimodal applications and media involving haptics. Second, they provide a shared cognitive grounding for design in different modalities. Past research has shown that users can associate stimuli, in different modalities, designed for shared perceptual or cognitive attributes (Hoggan and Brewster, 2007a). Facets encapsulate and structure users' cognitive schemas and are shared (or overlap considerably) among different modalities (e.g., emotion$_f$ attributes). Visual and auditory design already utilize a body of guidelines and cultural connotations established for affective design. Our results support haptic designers in creating sensations that are congruent with their visual and auditory analogs in their perceptual and affective attributes. The other, less-utilized senses in HCI (e.g., olfactory) can benefit by following a facet-based approach to design. In fact, the olfactory sense already has a basis for a faceted language. For example, Fragrantica, [1] an online perfume review website, presents perfumes according to their top, middle, and base notes as well as their longevity, sillage, usage context (e.g., day, summer), and user group (e.g., female > 25). These modalities can benefit from our procedure in developing their facets and characterizing their semantic structures for design.

We close by noting that rarely have the many challenges inherent in haptic evaluation (MacLean et al., 2017) been approached through the development of new, haptic-specific methodologies and evaluation instruments. Here, we offer a novel, scalable data collection approach to mapping users' comprehension of large sets of haptic signals; and report between- and within-subject data variation that can inform future instrument development.

Acknowledgments

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) (Grant no. NSERC RGPIN 239120-13) and the University of British Columbia's 4YF fellowship program. The user study was conducted under UBC Ethics certificate #H13-01646.

Appendix A. List of tags and their disagreement values

Tables A.7–A.10

[1] http://www.fragrantica.com/.

Table A.7
Sensation$_f$ tags and disagreement scores.

Index	Tag	Disagreement Score
1	short	0.08
2	smooth transition	0.09
3	irregular	0.11
4	pointy	0.11
5	ramping up	0.12
6	grainy	0.12
7	long	0.13
8	simple	0.17
9	firm	0.17
10	rough	0.17
11	wavy	0.17
12	continuous	0.17
13	discontinuous	0.17
14	bumpy	0.17
15	dynamic	0.2
16	regular	0.2
17	spiky	0.21
18	soft	0.22
19	springy	0.22
20	smooth	0.22
21	ramping down	0.24
22	complex	0.28
23	flat	0.28
24	ticklish	0.31

Table A.8
Emotion$_f$ tags and disagreement scores.

Index	Tag	Disagreement Score
1	rhythmic	0.14
2	attention-getting	0.16
3	agitating	0.18
4	unique	0.18
5	energetic	0.18
6	mechanical	0.19
7	familiar	0.2
8	surprising	0.21
9	urgent	0.22
10	natural	0.22
11	strange	0.23
12	predictable	0.24
13	uncomfortable	0.25
14	lively	0.25
15	calm	0.26
16	interesting	0.26
17	annoying	0.27
18	comfortable	0.27
19	pleasant	0.31
20	happy	0.31
21	angry	0.32
22	boring	0.32
23	creepy	0.32
24	sad	0.34
25	fear	0.36
26	funny	0.36

Table A.9
Metaphor$_f$ tags and disagreement scores.

Index	Tag	Disagreement Score
1	dancing	0.11
2	pulsing	0.11
3	getting close	0.11
4	cymbal	0.11
5	alarm	0.15
6	phone	0.15
7	morse code	0.16
8	heart beat	0.17
9	SOS	0.18
10	buzz	0.18
11	engine	0.19
12	sliding	0.2
13	tapping	0.21
14	game	0.22
15	going away	0.22
16	shaking	0.22
17	a door closing	0.22
18	stopping	0.22
19	growl	0.22
20	frogs	0.22
21	poking	0.23
22	coming or going	0.23
23	beep	0.24
24	horn	0.25
25	jumping	0.25
26	snoring	0.27
27	riding	0.28
28	clock	0.28
29	drums	0.28
30	breathing	0.3
31	electric shock	0.3
32	musical instruments	0.3
33	nature	0.31
34	bell	0.31
35	gun	0.31
36	pawing	0.31
37	celebration	0.31
38	walking	0.33
39	echo	0.33
40	explosion	0.33
41	chainsaw	0.33
42	animal	0.34
43	a spring	0.44
44	footsteps	0.44
45	a story	0.44

Table A.10
$Usage_f$ tags and disagreement scores.

Index	Tag	Disagreement Score
1	alarm	0.21
2	halfway	0.21
3	reminder	0.22
4	warning	0.22
5	running out of time	0.23
6	confirmation	0.23
7	speed up	0.24
8	overtime	0.24
9	slow down	0.25
10	interval/rep	0.25
11	above intended threshold	0.26
12	resume	0.26
13	one minute left	0.27
14	finish	0.27
15	incoming message	0.28
16	congratulations	0.28
17	get ready	0.3
18	milestone	0.3
19	encouragement	0.3
20	battery low	0.3
21	pause	0.3
22	warm up	0.31
23	cool down	0.31
24	below intended threshold	0.33
25	start	0.36

.

Appendix B. Tag removal summary

Table B.11

Table B.11
Percentage of tags removed by lay users. Each row represents the percentages of tags that are removed by at least x people (x=1 for ≥ 1 label) in each facet (columns).

Number of participants removing a tag	$Sensation_f$ (%)	$Emotion_f$ (%)	$Metaphor_f$ (%)	$Usage_f$ (%)
≥ 1	79	88	87	92
≥ 2	51	69	67	74
≥ 3	27	46	43	53
≥ 4	14	28	24	31
≥ 5	8	14	10	15
≥ 6	4	6	3	7
≥ 7	2	2	1	2
≥ 8	1	0	0	0
≥ 9	1	0	0	0

Appendix C. Rating correlations

Table C.12

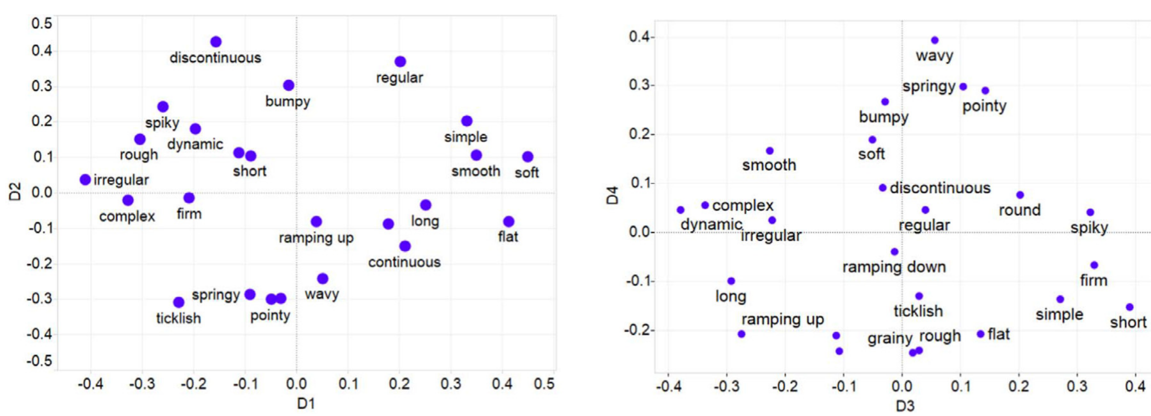
Table C.12
Results of Pearson correlation of the five rating scales. The correlation is applied on all participants' ratings for the 120 vibrations.

	Energy$_d$	Tempo$_d$	Roughness$_d$	Pleasantness$_d$	Arousal$_d$
Energy$_d$	1.00	0.48	0.74	-0.46	0.92
Tempo$_d$	0.48	1.00	0.52	-0.22	0.56
Roughness$_d$	0.74	0.52	1.00	-0.61	0.79
Pleasantness$_d$	-0.46	-0.22	-0.61	1.00	-0.53
Arousal$_d$	0.92	0.56	0.79	-0.53	1.00

.

Appendix D. Multidimensional scaling graphs on tag distances

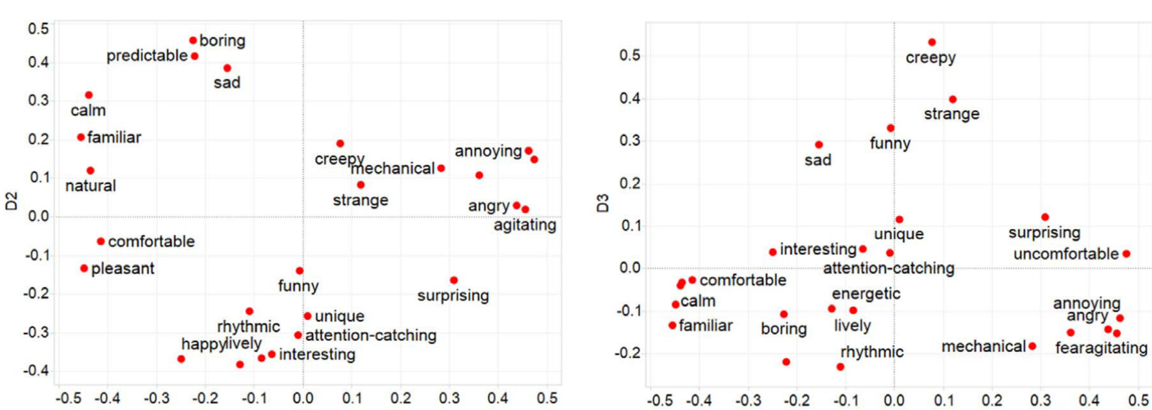
Figs. D.13–D.16



(a) Dimension 1 (simple/complex) *vs.* 2 (discontinuous/continuous)

(b) Dimension 3 (short/long) *vs.* 4 (rough/smooth)

Fig. D.13. Spatial configuration of the tags for the sensation$_f$ facet confirms the four identified dimensions. Specifically, contrasting tags according to each dimension are well-separated, and the semantically-related tags are close together along each dimension (supporting convergent and discriminant validity). For example, simple, regular, flat, smooth and soft are far from irregular, complex, and rough.



(a) Dimension 1 (comfortable/agitating) vs. 2 (boring/lively)

(b) Dimension 1 (comfortable/agitating) vs. 3 (strange/predictable)

Fig. D.14. Spatial configuration of the tags for the emotion$_f$ facet confirms the three identified dimensions and supports convergent and discriminant validity.

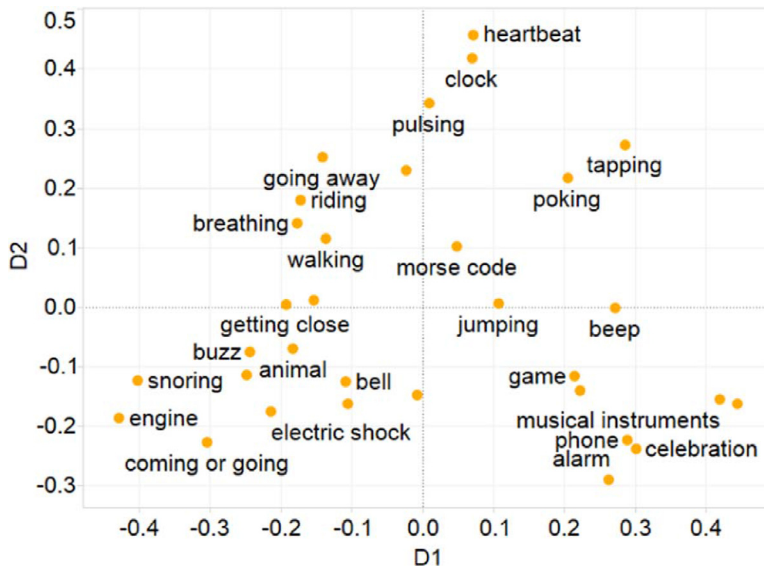


Fig. D.15. Spatial configurations of tags for the metaphor$_f$ facet. Dimension 1 (on-off—ongoing$_d$) vs. dimension 2 (natural—mechanical$_d$). Semantically-related tags, according to a dimension, are close along the dimension (e.g., drums, celebration, alarm) and contrasting tags are far from each other (e.g., heartbeat vs. engine or alarm). This definition leaves a few tags, such as clock (among the natural, calm sensations) and snoring (with mechanical, annoying and ongoing tags).

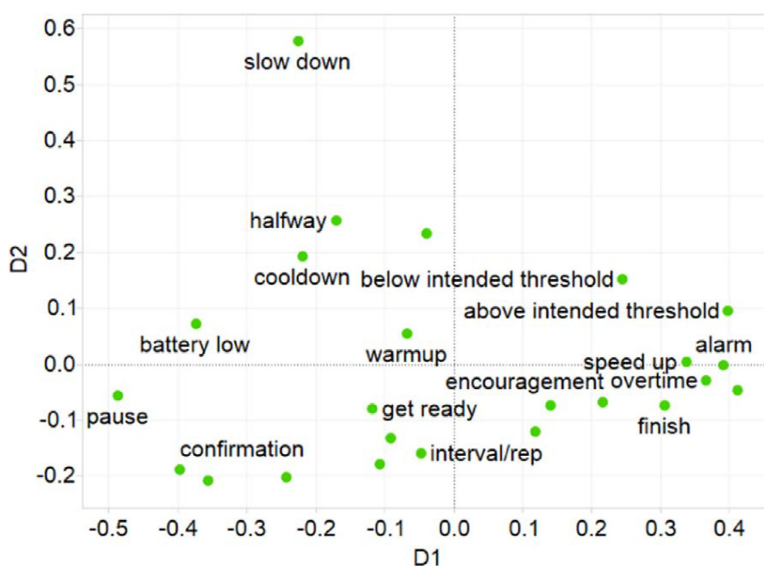


Fig. D.16. Spatial configurations of tags for the usage$_f$ facets. Dimension 1 (urgent- awareness notifications). Dimension 2 is not used in our analysis. Along Dimension 1, tags have increasing urgency and attention demand from left to right, supporting convergent and discriminant validity for the semantics of the dimension.

.

Appendix E. Individual differences in vibrations

Fig. E.17

	Energy	Tempo	Roughness	Valence	Arousal	Sensation	Emotion	Metaphor	Usage
v-09-10-3-56	0.44	0.62	0.59	0.91	0.79	0.02	0.39	0.33	0.29
v-09-10-4-25	0.35	0.59	1.26	0.96	0.81	0.3	0.43	0.07	0.28
v-09-10-6-46	0.4	0.4	0.67	0.91	0.84	0.27	0.2	0.44	0.37
v-10-28-7-35	0.99	1.49	1.16	1.19	0.53	0.16	0.31	0.41	0.4
v-09-12-1-0	0.46	1.39	0.86	1.07	0.83	0.27	0.41	0.24	0.25
v-09-09-8-11	1.08	0.54	0.59	0.63	0.5	0.15	0.22	0.22	0.27
v-09-09-8-20	0.92	0.65	1.1	0.91	0.76	0.22	0.17	0.19	0.3
v-09-09-8-20-cpy	0.85	0.36	0.85	0.76	0.69	0.22	0.14	0.19	0.21
v-09-09-8-24	0.89	1.1	1.11	1.38	0.62	0.2	0.36	0	0.33
v-09-10-11-55	0.47	0.96	1.1	1.01	0.83	0.09	0.17	0.22	0.23
v-09-10-12-11	0.57	0.53	0.28	0.99	0.59	0.15	0.09	0.11	0.06
v-09-10-12-13	0.28	0.58	0.65	1.18	0.38	0.09	0.26	0.14	0.26
v-09-10-12-16	0.4	1.11	0.79	1.33	0.44	0.17	0.27	0.44	0.24
v-09-10-12-2	0.79	0.36	0.52	0.86	0.28	0.16	0.2	0.22	0.22
v-09-10-12-6	0.77	0.62	0.46	0.93	1.01	0.03	0.17	0.16	0.27
v-09-10-12-9	0.53	0.5	0.66	0.88	0.18	0.21	0.26	0.27	0.28
v-09-10-12-9-cpy	0.68	0.33	0.73	1.11	0.66	0.24	0.19	0.38	0.39
v-09-10-3-52	0.14	0.26	0.57	0.83	0.64	0.2	0.28	0.24	0.22
v-09-10-4-2	0.6	0	0.71	1.26	0.66	0.25	0.37	0.29	0.32
v-09-10-4-20	0.91	1.21	0.98	1.49	0.66	0.37	0.29	0.27	0.37
v-09-10-4-23	0.49	0.59	1.11	1.11	0.79	0.16	0.25	0.31	0.31
v-09-10-4-6	0.72	0.63	0.88	1.02	0.98	0.07	0.15	0.15	0.13
v-09-10-6-16	0.86	0.35	0.59	1.19	0.91	0.3	0.3	0.26	0.4
v-09-10-6-22	0.47	0.72	0.71	0.93	1.16	0.26	0.29	0.17	0.11
v-09-10-6-27	0.14	0	0.78	0.92	0.14	0.13	0.22	0.18	0.13
v-09-10-6-38	0.35	0.68	0.99	1.07	0.33	0.13	0.4	0.26	0.35
v-09-10-6-43	0.53	0.82	0.83	0.88	0.71	0.19	0.18	0.32	0.29
v-09-10-6-5	0.2	0.35	0.44	0.94	0.74	0.13	0.38	0.2	0.33
v-09-10-6-59	0.64	1.02	0.63	0.64	0.53	0.14	0.19	0.22	0.22
v-09-10-7-34	1.09	1.04	1.1	1.33	0.89	0.11	0.3	0.19	0.31
v-09-10-7-36	0.53	0.66	1.02	0.88	0.46	0.17	0.25	0.11	0.26
v-09-10-7-9	0.45	0.53	1.05	1.08	1.11	0.18	0.21	0.22	0.19
v-09-10-8-5	0.53	0.69	1.14	0.84	0.91	0.22	0.39	0.33	0.38
v-09-10-8-7	1.01	0.47	0.65	0.8	0.77	0.13	0.04	0.04	0.24
v-09-10-8-7-cpy	0.72	0.47	0.66	0.9	0.92	0.11	0.17	0.07	0.2
v-09-11-3-12	0.69	0.53	0.88	1.29	0.55	0.3	0.31	0.11	0.33
v-09-11-3-16	0.75	0.28	0.82	0.93	0.63	0.02	0.24	0.08	0.19
v-09-11-3-19	0.4	0.78	1.06	1.27	0.17	0.28	0.29	0.3	0.26
v-09-11-3-21	0.49	0.49	0.4	0.62	0.67	0.13	0.33	0.38	0.3
v-09-11-3-24	0.89	0.4	0.89	0.62	0.59	0.2	0.21	0.41	0.27
v-09-11-3-4	0.5	0.5	0.44	0.47	0.42	0.24	0.26	0.28	0.3
v-09-11-3-43	0.5	0.5	0.74	1.06	0.58	0.35	0.31	0.29	0.44
v-09-11-3-50	1.02	0.99	0.48	0.99	1.01	0.22	0.36	0.22	0.36
v-09-11-3-54	0.72	0.56	1.11	0.63	0.58	0.22	0.12	0.25	0.18
v-09-11-3-56	0.89	0.59	0.89	1.04	0.96	0.13	0.25	0.47	0.36
v-09-11-3-8	0.42	0.33	0.56	0.85	0.58	0.18	0.33	0.42	0.27
v-09-11-4-1	0.62	0.49	0.89	0.44	0.44	0.24	0.3	0.25	0.38
v-09-11-4-12	0.88	0.42	0.44	0.58	0.94	0.24	0.33	0.33	0.22
v-09-11-4-22	0.2	1.23	0.52	0.81	0.67	0.21	0.25	0.41	0.33
v-09-11-4-3	0.31	0.5	1.11	0.67	0.67	0.09	0.14	0.29	0.24
v-09-11-4-41	1.14	0.79	1.42	0.93	0.76	0.32	0.35	0.35	0.36
v-09-11-4-41-cpy	1.11	0.99	1.54	1.22	1.21	0.37	0.3	0.37	0.42
v-09-11-4-54	1.19	0.46	0.76	1.36	0.17	0.2	0.22	0.21	0.2
v-09-11-4-8	1.22	0.45	0.68	1.31	0.74	0.19	0.21	0	0.46
v-09-12-1-19	0.69	0.49	0.74	0.74	0.67	0.33	0.29	0.49	0.33
v-09-12-1-23	0.71	0.65	1.14	1.17	1.17	0.13	0.11	0.07	0.19
v-09-12-1-29	0.89	0.85	1.23	0.9	1.03	0.17	0.27	0.15	0.17
v-09-12-1-39	1.29	1.52	1.7	1.22	1.07	0.2	0.17	0.38	0.35
v-09-12-1-48	0.5	0.39	0.65	1.14	0.43	0.19	0.38	0.28	0.23
v-09-12-1-53	0.81	0.44	1.01	1.11	0.81	0.15	0.24	0.53	0.32
v-09-12-2-17	0.81	1.11	1.43	1.19	1.26	0.17	0.17	0.31	0.35
v-09-12-2-20	0.54	0.83	1.17	1.24	1.11	0.14	0.32	0.17	0.2
v-09-12-2-23	0.92	0.58	1.02	0.89	0.75	0.13	0.3	0.16	0.23
v-09-12-2-40	0.58	0.89	0.45	1.22	1.4	0.2	0.42	0.41	0.48
v-09-12-8-10	0	1.22	1.07	1.98	0.53	0.22	0.37	0.28	0.38

	Energy	Tempo	Roughness	Valence	Arousal	Sensation	Emotion	Metaphor	Usage
v-09-12-8-13	1.18	1.04	1.14	0.8	1.15	0.04	0.25	0.14	0.32
v-09-12-8-21	0.93	1.03	1.01	0.86	0.84	0.09	0.11	0.17	0.19
v-09-12-8-27	0.62	1.19	1.26	0.96	0.69	0.06	0.13	0.33	0.39
v-09-12-8-30	0.17	1.21	0.93	1.11	0.6	0.2	0.41	0.41	0.39
v-09-12-8-32	0.74	0.57	0.96	1.53	1.04	0.27	0.3	0.44	0.34
v-09-16-1-43	0.67	0.62	0.96	1.04	0.69	0.26	0.28	0.22	0.28
v-09-16-1-43-cpy	0.69	1.01	1.19	1.04	0.62	0.24	0.33	0.33	0.24
v-09-16-1-56	0.71	0.5	0.91	0.8	0.53	0.17	0.23	0.13	0.19
v-09-18-1-55	0.56	0.75	1	1.25	0.5	0.27	0.27	0.11	0.32
v-09-18-2-7	0.46	0.71	0.81	1.36	0.3	0.13	0.23	0.28	0.37
v-09-18-4-12	1.22	0.83	0.79	0.66	0.66	0.16	0.26	0.25	0.38
v-09-18-4-15	0.94	0.2	0.59	0.84	0.52	0.14	0.26	0.26	0.33
v-09-18-4-16	1.12	0.63	0.74	1.21	0.96	0.17	0.43	0.24	0.32
v-09-18-4-18	0.54	0.59	0.77	1.05	0.57	0.06	0.18	0.14	0.15
v-09-18-4-22	1.43	1.11	1.13	0.68	0.83	0.23	0.31	0.28	0.33
v-09-18-4-56	1.75	1.62	0.94	1.29	1.36	0.32	0.39	0.38	0.26
v-09-23-6-24	1.7	0.96	1.41	1.26	0.89	0.22	0.29	0.3	0.35
v-09-23-6-24-cpy	1.33	0.99	1.16	1.58	0.59	0.28	0.25	0.26	0.4
v-09-26-1-39	1.53	1.23	0.99	1.23	0.91	0.14	0.19	0.11	0.17
v-10-09-1-1	0.33	0.73	0.98	1.09	0.88	0.24	0.36	0.26	0.44
v-10-09-1-11	0.44	0.89	0.67	0.74	0.89	0.19	0.19	0.44	0.36
v-10-09-1-12	1	0.82	0.58	0.92	1.31	0.15	0.33	0.22	0.22
v-10-09-1-14	0.9	0.52	0.54	0.93	0.66	0.07	0.05	0.08	0.27
v-10-09-1-16	0.75	0.83	0.72	0.77	0.63	0.15	0.22	0	0.28
v-10-09-1-16-cpy	0.73	0.76	1.04	1.02	0.54	0.15	0.25	0	0.24
v-10-09-1-20	0.83	0.78	0.67	1.26	0.61	0.17	0.35	0.4	0.36
v-10-09-1-8	1.39	1.11	0.88	1.25	1.17	0.3	0.33	0.22	0.24
v-10-09-1-8-cpy	1.17	1.22	0.72	1.32	1.08	0.21	0.37	0.33	0.28
v-10-09-5-0	0.4	0.2	0.79	0.81	0.81	0.22	0.17	0.26	0.26
v-10-09-5-2	0.69	0.46	0.45	1.42	0.66	0.22	0.22	0.29	0.34
v-10-09-5-4	0.66	0.96	0.69	1.11	0.45	0.36	0.35	0.38	0.31
v-10-09-5-7	0.44	0.28	0.58	1.17	0.49	0.24	0.24	0.25	0.27
v-10-09-5-7-cpy	0.28	0.28	0.44	0.92	0.49	0.22	0.2	0.36	0.32
v-10-10-1-10	0.6	0.45	0.51	0.66	0.99	0.31	0.26	0.44	0.36
v-10-10-1-10-cpy	1.12	1.06	0.84	1.19	0.45	0.33	0.36	0.5	0.33
v-10-10-1-18	0.88	0.48	1.39	1.22	0.58	0.33	0.35	0.29	0.38
v-10-10-1-21	0.43	0.67	0.39	1.05	0.47	0.11	0.17	0.11	0.14
v-10-10-1-5	0.59	0.8	0.93	1.11	0.63	0.11	0.31	0.06	0.18
v-10-18-11-11	0.15	0.38	0.88	0.69	0.15	0.24	0.19	0.14	0.34
v-10-21-2-48	0.85	0.58	0.65	1.15	0.66	0.11	0.21	0.2	0.15
v-10-21-3-11	0.14	0.28	0.47	0.99	0.28	0.13	0.24	0.33	0.24
v-10-21-3-17	0.15	0.67	0.68	0.92	0.15	0.14	0.28	0.19	0.3
v-10-21-3-2	0.74	0.4	1.19	0.67	0.59	0.16	0.33	0.48	0.28
v-10-21-3-21	0.89	1.31	0.55	0.78	1.12	0.25	0.33	0.46	0.32
v-10-21-3-30	1.17	0.67	1.06	1.22	0.67	0.11	0.39	0.29	0.37
v-10-21-3-33	1.35	0.15	1.33	1.75	0.96	0.13	0.35	0.17	0.15
v-10-21-3-39	0.89	0.43	0.53	0.58	0.47	0.04	0.21	0.08	0.24
v-10-21-3-4	0.64	0.26	1.43	1.08	0.47	0.04	0.17	0.12	0.16
v-10-21-3-45	0.14	0.28	0.39	0.8	0.28	0.14	0.13	0.13	0.13
v-10-21-3-45-cpy	0.47	0.31	0.57	0.67	0.36	0.22	0.18	0.17	0.1
v-10-21-3-7	1.08	0.38	1.03	1.03	0.92	0.19	0.24	0.19	0.23
v-10-23-1-10	0	0.4	0.59	0.86	0.52	0.2	0.32	0.22	0.17
v-10-23-1-16	0.59	0.22	0.44	1.11	0.35	0.2	0.24	0.22	0.15
v-10-23-1-21	1.27	0.69	0.85	0.86	0.91	0.17	0.24	0.2	0.17
v-10-23-1-23	0.81	0.69	0.81	0.98	0.53	0.22	0.41	0.27	0.36
v-10-23-1-24	0.57	0.58	0.58	0.69	0.65	0.06	0.16	0.3	0.28
v-10-28-7-22	0	0.59	1.04	1.38	0.44	0.15	0.21	0.22	0.33
v-10-28-7-22-cpy	0	0.72	0.69	1.33	0.44	0.11	0.21	0.17	0.32
v-10-28-7-23	1.15	0.62	0.75	0.82	0.53	0	0.07	0.06	0.14
v-10-28-7-26	1.17	0.42	0.5	0.83	0.67	0.11	0.31	0.26	0.38
v-10-28-7-29	0.81	0.74	1.01	1.14	1.06	0.29	0.28	0.44	0.39
v-10-28-7-31	0.74	0.69	0.79	1.33	0.94	0.11	0.21	0.27	0.36
v-10-28-7-33	1.13	0.58	1.06	0.42	0.99	0.18	0.35	0.25	0.33
v-10-28-7-36	1.56	0.67	0.42	0.96	0.67	0.17	0.27	0.31	0.31
v-10-29-4-20	0.81	1.19	0.89	1.63	1.26	0.16	0.4	0.19	0.44
v-10-29-4-22	0.81	0.62	1.04	0.62	0.72	0.17	0.22	0.22	0.43

Fig. E.17. Vibration disagreement scores for the five rating scales and the four facets. High color saturation denotes high disagreement scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Appendix F. Between-facet tag linkages

Figs. F.18–F.20

	bumpy	complex	cont.	discont.	dynamic	firm	flat	grainy	irregular	long	pointy	rampdwn	rampup	regular	rough	short	simple	smooth	soft	spiky	springy	ticklish	wavy
agitating	0.16	0.36	0.16	0.46	0.39	0.21	0.04	0.29	0.36	0.2	0	0	0.23	0.3	0.67	0.3	0.08	0.06	0	0.32	0	0.24	0
angry	0.11	0.1	0.14	0.26	0.25	0.21	0	0.16	0.17	0.15	0	0.07	0.19	0.19	0.5	0.19	0.04	0	0	0.26	0	0.09	0
annoying	0.16	0.24	0.34	0.35	0.42	0.22	0.09	0.39	0.25	0.21	0	0.09	0.37	0.24	0.66	0.31	0.11	0.03	0.03	0.22	0	0.15	0
boring	0.14	0	0.37	0.13	0.14	0.14	0.47	0.36	0.04	0.24	0	0.2	0.09	0.24	0.11	0.19	0.39	0.28	0.27	0.05	0	0	0.1
calm	0.23	0.06	0.22	0.39	0.15	0.08	0.43	0.2	0.09	0.32	0.05	0.27	0.13	0.48	0.05	0.25	0.54	0.62	0.58	0.1	0	0	0.05
comfortable	0.5	0.21	0.26	0.52	0.37	0.03	0.22	0.26	0.27	0.21	0.04	0.19	0.16	0.4	0.09	0.26	0.4	0.55	0.56	0.08	0.04	0	0.07
creepy	0	0.13	0.13	0.02	0.07	0	0.12	0.11	0.06	0.07	0	0	0.13	0	0.05	0.07	0.05	0	0	0	0	0.2	0
energetic	0.44	0.1	0	0.29	0.11	0.08	0	0.13	0.04	0.1	0	0	0.1	0.31	0.04	0.2	0.26	0	0.1	0.28	0.13	0	0
familiar	0.29	0	0.19	0.31	0.05	0.1	0.2	0.17	0	0.31	0	0.2	0.08	0.45	0.06	0.26	0.6	0.37	0.48	0.12	0	0	0
fear	0.04	0.16	0.11	0.18	0.21	0.17	0.08	0.23	0.29	0.17	0	0	0.16	0.12	0.27	0.22	0.04	0	0	0.12	0	0.35	0
funny	0.08	0.3	0.06	0.12	0.19	0	0	0.15	0.21	0	0	0	0.06	0	0.13	0.12	0.04	0.05	0.06	0.14	0.22	0.15	0.2
happy	0.41	0.18	0.05	0.31	0.21	0	0.06	0.2	0.12	0.14	0	0	0.05	0.34	0.1	0.09	0.21	0.2	0.22	0.15	0.1	0	0.1
interesting	0.48	0.46	0.14	0.45	0.46	0.05	0	0.16	0.33	0.15	0	0	0.21	0.23	0.2	0.21	0.23	0.22	0.14	0.23	0.06	0.06	0
lively	0.54	0.33	0.03	0.58	0.39	0.12	0	0.16	0.34	0.18	0	0.04	0.18	0.41	0.2	0.24	0.3	0.22	0.18	0.32	0.05	0.04	0.05
mechanical	0.24	0.32	0.36	0.58	0.51	0.2	0.1	0.44	0.4	0.26	0.03	0.17	0.4	0.41	0.61	0.28	0.28	0.16	0.08	0.24	0	0.08	0
natural	0.3	0.09	0.09	0.27	0.13	0.06	0.13	0.12	0.08	0.14	0	0.13	0	0.34	0	0.18	0.34	0.46	0.43	0.1	0	0	0.09
pleasant	0.51	0.18	0.11	0.38	0.18	0	0.14	0.25	0.1	0.15	0	0.05	0.07	0.41	0.03	0.25	0.49	0.44	0.56	0.15	0.06	0	0.12
predictable	0.22	0.03	0.41	0.23	0.22	0.09	0.36	0.37	0.06	0.32	0	0.13	0.24	0.44	0.19	0.14	0.5	0.4	0.37	0.07	0	0	0.06
rhythmic	0.55	0.34	0.03	0.54	0.29	0.2	0.04	0.08	0.34	0.25	0.05	0.08	0.03	0.43	0.23	0.12	0.28	0.22	0.21	0.26	0.05	0.04	0
sad	0.04	0.06	0.24	0.08	0.09	0.11	0.5	0.05	0	0.31	0	0.4	0.18	0.15	0	0.06	0.26	0.3	0.24	0	0	0	0
strange	0.22	0.39	0.28	0.33	0.43	0.05	0.14	0.28	0.45	0.14	0	0.23	0.28	0.11	0.23	0.25	0.2	0.16	0.1	0.15	0	0.05	0
surprising	0.13	0.48	0.26	0.29	0.47	0.19	0	0.15	0.43	0.04	0	0.18	0.35	0.03	0.33	0.17	0.03	0.11	0.04	0.14	0.09	0.15	0
uncomfortable	0	0.29	0.21	0.27	0.28	0.35	0.06	0.25	0.15	0.21	0	0.17	0.42	0.13	0.45	0.25	0.1	0.04	0	0.14	0	0.21	0
unique	0.28	0.68	0.18	0.41	0.55	0.05	0.05	0.16	0.52	0.18	0	0.09	0.14	0.14	0.23	0.14	0.11	0.22	0.11	0.27	0.06	0.11	0
urgent	0.19	0.37	0.29	0.49	0.46	0.25	0.04	0.34	0.35	0.23	0	0.11	0.34	0.31	0.71	0.26	0.17	0.08	0	0.27	0	0.16	0

Figure F.18: Co-occurrence of $\mathsf{sensation}_f$ and $\mathsf{emotion}_f$ tags

Fig. F.18. Co-occurrence of sensation_f and emotion_f tags.

	bumpy	complex	cont.	discont.	dynamic	firm	flat	grainy	irregular	long	pointy	rampdwn	rampup	regular	rough	short	simple	smooth	soft	spiky	springy	ticklish	wavy
phone	0.18	0.24	0.1	0.28	0.03	0.29	0.07	0.04	0.13	0.2	0	0.14	0.14	0.16	0.39	0.19	0.18	0.08	0	0.26	0	0.09	0
walking	0.05	0.07	0	0.06	0.16	0	0.27	0.06	0.06	0.07	0	0.13	0	0.07	0	0	0.1	0.06	0.07	0.08	0	0	0
jumping	0.13	0.19	0	0.14	0	0	0	0.05	0.11	0.06	0	0	0.19	0.06	0.09	0.06	0	0.05	0.06	0.07	0.25	0	0
going away	0.1	0	0	0.02	0.03	0	0	0	0	0.15	0	0.29	0	0	0	0	0	0	0	0	0	0	0
breathing	0	0	0	0.02	0.16	0	0	0.06	0	0.08	0	0.29	0.07	0.07	0	0	0.05	0.06	0.07	0	0	0	0
horn	0.04	0.13	0.13	0.1	0.11	0.33	0.11	0.05	0.16	0.13	0	0.11	0.19	0.06	0.17	0.06	0.04	0.05	0	0.07	0	0.17	0
poking	0.22	0.1	0.1	0.23	0.28	0.16	0.08	0.09	0.14	0	0	0	0.05	0.17	0.11	0.46	0.26	0.09	0.2	0.23	0	0	0
beep	0.23	0.22	0.19	0.37	0.15	0.2	0.15	0.13	0.24	0.15	0	0.1	0.04	0.28	0.32	0.44	0.26	0.2	0.18	0.2	0	0.12	0.06
pulsing	0.24	0	0	0.18	0.12	0	0.09	0.14	0	0.06	0	0.09	0.06	0.24	0.12	0.17	0.2	0.19	0.39	0	0	0	0.17
snoring	0.12	0	0.28	0.08	0.25	0.09	0.35	0.23	0	0.29	0	0.17	0.28	0.15	0.16	0.06	0.2	0.23	0.16	0	0	0	0.15
sliding	0.16	0.18	0.29	0.06	0.09	0	0	0.05	0	0.12	0	0	0.29	0.03	0.08	0	0	0.2	0.11	0	0	0.14	0.36
nature	0.08	0	0.12	0.08	0.07	0	0.3	0.15	0.11	0.06	0	0	0.06	0.09	0.04	0.06	0.21	0.15	0.18	0.14	0	0	0
morse code	0.13	0.13	0	0.11	0.13	0.13	0	0	0.06	0	0	0	0	0.1	0	0	0.05	0.05	0	0.08	0.33	0.2	0
riding	0.21	0.06	0.06	0.06	0.11	0	0	0.15	0.11	0.13	0	0.21	0.06	0	0.04	0	0	0	0.06	0	0.25	0	0
buzz	0.11	0.1	0.35	0.11	0.27	0	0.22	0.21	0.13	0.31	0	0.15	0.25	0.25	0.19	0.2	0.3	0.17	0.29	0	0	0.1	0
animal	0.15	0.12	0.35	0.24	0.43	0	0.16	0.28	0.14	0.32	0	0.05	0.35	0.31	0.22	0.2	0.31	0.24	0.27	0.17	0	0.06	0.07
engine	0.19	0.29	0.49	0.14	0.03	0.06	0.06	0.43	0.22	0.17	0	0.17	0.53	0.12	0.22	0	0.16	0.18	0.04	0.09	0	0.07	0
clock	0.13	0	0.06	0.1	0.22	0.24	0.22	0.05	0	0.07	0	0	0	0.13	0	0.13	0.18	0.16	0.19	0.15	0	0	0
musical instrume	0.14	0.24	0.05	0.26	0.06	0.21	0	0.08	0.21	0.15	0	0.07	0.05	0.19	0.36	0.14	0.04	0.08	0.05	0.26	0	0	0
pawing	0.09	0.06	0.06	0.12	0.24	0	0	0	0.05	0.06	0	0	0	0.09	0.04	0.13	0.04	0.21	0.18	0.07	0	0	0.22
game	0.32	0.26	0.11	0.42	0.22	0.15	0	0.13	0.34	0.04	0	0.05	0.11	0.26	0.27	0.38	0.21	0.17	0.15	0.12	0	0.12	0.07
drums	0.14	0.28	0.05	0.28	0.09	0.21	0	0.08	0.25	0.24	0.11	0.07	0.05	0.19	0.28	0.14	0.04	0.12	0.09	0.21	0	0	0
gun	0.08	0.06	0.12	0.12	0.03	0.11	0	0.3	0.11	0	0	0	0.12	0.06	0.21	0.3	0.04	0	0.06	0.21	0	0.46	0
getting close	0.05	0.07	0.15	0	0.03	0	0	0	0	0	0	0	0.07	0	0	0	0.05	0	0	0	0	0	0
bell	0.04	0	0.18	0.06	0.4	0	0.2	0.05	0.05	0.06	0	0.2	0.12	0.09	0.13	0.12	0.26	0.15	0.12	0.07	0	0	0
alarm	0.23	0.28	0.19	0.48	0.11	0.2	0.04	0.11	0.35	0.22	0	0.16	0.22	0.29	0.56	0.16	0.18	0.11	0.03	0.23	0	0.09	0.05
heartbeat	0.31	0	0	0.31	0.1	0.14	0	0.04	0	0.29	0	0.2	0.05	0.29	0.04	0.19	0.32	0.24	0.32	0.26	0	0	0
SOS	0	0.2	0	0.11	0.38	0.13	0	0	0.11	0.14	0	0	0.07	0.03	0.05	0.13	0	0	0	0.23	0	0	0
tapping	0.52	0.17	0.03	0.61	0.3	0.18	0.07	0.21	0.19	0.23	0.04	0.04	0.09	0.57	0.29	0.26	0.38	0.18	0.31	0.36	0.04	0	0.04
coming or going	0.11	0.19	0.29	0.15	0.06	0.14	0.07	0.24	0.09	0.2	0	0.28	0.57	0.14	0.18	0.1	0.07	0.12	0	0.05	0	0.18	0
celebration	0.04	0.13	0	0.1	0.03	0	0	0	0	0.2	0	0	0.13	0.1	0.13	0.06	0	0.11	0.06	0.07	0	0	0
electric shock	0.04	0	0.18	0.1		0.2	0.1	0.29	0.1	0.12	0	0	0	0.09	0.29	0.24	0.13	0	0.06	0	0	0.29	0

Fig. F.19. Co-occurrence of sensation$_f$ and metaphor$_f$ tags.

	bumpy	complex	cont.	discont.	dynamic	firm	flat	grainy	irregular	long	pointy	rampdwn	rampup	regular	rough	short	simple	smooth	soft	spiky	springy	ticklish	wavy
encouragement	0.44	0.42	0.16	0.47	0.45	0.04	0	0.2	0.36	0.1	0	0.08	0.19	0.34	0.26	0.19	0.18	0.29	0.22	0.21	0.05	0.1	0.05
reminder	0.53	0.24	0.2	0.53	0.32	0.11	0.07	0.16	0.22	0.15	0	0.07	0.09	0.46	0.12	0.35	0.41	0.35	0.46	0.28	0.05	0	0.09
congratulations	0.05	0.07	0.07	0	0.03	0	0.14	0	0.06	0.07	0	0.14	0	0.03	0	0	0.05	0.12	0.07	0	0	0	0
cooldown	0.36	0.17	0.13	0.27	0.21	0	0.12	0.19	0.16	0.22	0.09	0.12	0.17	0.25	0.17	0.13	0.23	0.11	0.21	0.1	0.09	0	0.09
running out of time	0.46	0.28	0.2	0.61	0.52	0.15	0	0.32	0.33	0.18	0	0.09	0.3	0.5	0.49	0.2	0.19	0.25	0.24	0.21	0.04	0.1	0.04
pause	0.13	0.08	0.16	0.3	0.13	0.12	0.23	0.18	0.11	0.13	0.08	0.06	0.08	0.25	0.06	0.42	0.45	0.29	0.44	0.23	0	0	0
milestone	0.54	0.24	0.33	0.46	0.33	0.08	0.08	0.19	0.23	0.06	0.05	0.04	0.09	0.34	0.18	0.48	0.4	0.25	0.24	0.29	0.05	0	0.05
speed up	0.39	0.25	0.14	0.59	0.56	0.1	0	0.25	0.31	0.16	0	0.03	0.41	0.49	0.34	0.16	0.16	0.28	0.16	0.35	0.04	0.15	0
confirmation	0.17	0.05	0.18	0.2	0.11	0.14	0.2	0.08	0.08	0	0	0.07	0.05	0.16	0.07	0.47	0.42	0.24	0.31	0.21	0	0	0
halfway	0.2	0.06	0.16	0.12	0.09	0	0	0.09	0.05	0.06	0	0	0.11	0.17	0.04	0.06	0.2	0.28	0.16	0.19	0	0	0.15
above threshold	0.39	0.28	0.08	0.58	0.43	0.17	0	0.15	0.26	0.22	0	0.07	0.27	0.44	0.44	0.28	0.21	0.35	0.22	0.26	0.04	0.12	0.04
slow down	0.14	0.1	0.19	0.17	0.17	0	0.29	0.08	0	0.29	0	0.21	0.29	0.27	0.04	0.05	0.18	0.29	0.28	0	0	0	0
interval/rep	0.28	0.14	0.17	0.38	0.18	0.14	0.23	0.16	0.13	0.11	0.06	0.14	0.14	0.33	0.31	0.39	0.37	0.19	0.31	0.15	0.06	0.05	0.06
warmup	0.36	0.12	0.15	0.32	0.22	0.05	0.1	0.27	0.18	0.12	0	0	0.23	0.31	0.27	0.19	0.36	0.14	0.19	0.17	0.07	0	0.07
incoming msg	0.31	0.16	0.31	0.26	0.22	0	0.22	0.25	0.18	0.24	0	0	0.08	0.22	0.13	0.36	0.34	0.14	0.27	0.13	0	0	0.07
one minute left	0.46	0.24	0.2	0.5	0.38	0.07	0.04	0.32	0.27	0.09	0	0.11	0.14	0.42	0.32	0.35	0.34	0.21	0.23	0.22	0.05	0.04	0
finish	0.3	0.34	0.3	0.43	0.33	0.24	0.12	0.33	0.23	0.31	0	0.08	0.15	0.35	0.66	0.31	0.3	0.08	0.09	0.16	0.05	0.09	0
resume	0.29	0.07	0.25	0.34	0.19	0.1	0.1	0.16	0.14	0.04	0	0.05	0.15	0.25	0.15	0.48	0.44	0.26	0.32	0.2	0	0	0.06
alarm	0.34	0.33	0.24	0.59	0.54	0.14	0.03	0.31	0.34	0.31	0	0.06	0.38	0.44	0.59	0.26	0.22	0.18	0.14	0.25	0	0.13	0
get ready	0.46	0.29	0.29	0.51	0.36	0.1	0.16	0.24	0.17	0.34	0	0.16	0.21	0.5	0.13	0.29	0.49	0.27	0.36	0.31	0.04	0.07	0.04
start	0.28	0.11	0.11	0.16	0.18	0	0	0.05	0.05	0.06	0	0	0.11	0.15	0.04	0.06	0.12	0.1	0.11	0.13	0.18	0	0
battery low	0.23	0.04	0.21	0.35	0.26	0.15	0.19	0.23	0.13	0.11	0	0.19	0.14	0.34	0.12	0.36	0.38	0.32	0.35	0.2	0	0	0
warning	0.5	0.22	0.27	0.6	0.45	0.24	0.08	0.33	0.26	0.16	0	0.08	0.24	0.57	0.31	0.36	0.33	0.27	0.29	0.24	0.03	0.09	0.06
overtime	0.36	0.16	0.24	0.52	0.4	0.2	0.07	0.28	0.28	0.33	0	0.1	0.27	0.45	0.6	0.25	0.28	0.15	0.24	0.2	0	0.08	0.04
below threshold	0.46	0.15	0.12	0.53	0.33	0.11	0.07	0.14	0.14	0.39	0	0.19	0.32	0.5	0.32	0.21	0.3	0.32	0.29	0.19	0.05	0.04	0

Fig. F.20. Co-occurrence of sensation$_f$ and usage$_f$ tags.

References

Brown, L.M., Brewster, S.A., Purchase, H.C., 2006. Tactile crescendos and sforzandos: applying musical techniques to tactile icon design. In: Proceedings of Extended Abstracts on Human Factors in Computing Systems (CHI EA), pp. 610–615.

Brunet, L., Megard, C., Panëels, S., Changeon, G., Lozada, J., Daniel, M.P., Darses, F., 2013. Invitation to the voyage: the design of tactile metaphors to fulfill occasional travelers' needs in transportation networks. In: Proceedings of the IEEE World Haptics Conference (WHC), pp. 259–264.

Carbon, C.C., Jakesch, M., 2013. A model for haptic aesthetic processing and its implications for design. Proceedings of the IEEE 101 (9), 2123–2133.

Chan, A., MacLean, K., McGrenere, J., 2008. Designing haptic icons to support collaborative turn-taking. Int. J. Hum.-Comput. Stud. 66 (5), 333–355.

Chen, X., Shao, F., Barnes, C., Childs, T., Henson, B., 2009. Exploring relationships between touch perception and surface physical properties. Int. J. Des. 3 (2), 67–76.

Cox, T.F., Cox, M.A., 2000. Multidimensional Scaling. CRC Press.

Curran, P.G., 2016. Methods for the detection of carelessly invalid responses in survey data. J. Exp. Soc. Psychol., 66, pp. 4–19, Rigorous and Replicable Methods in Social Psychology. URL ⟨http://www.sciencedirect.com/science/article/pii/S0022103115000931⟩.

Doizaki, R., Watanabe, J., Sakamoto, M., 2014. A system for evaluating tactile feelings expressed by sound symbolic words. In: Auvray, M., Duriez, C., (Eds.), Haptics: Neuroscience, Devices, Modeling, and Applications: Proceedings of Eurohaptics, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 32–39. ⟨https://dx.doi.org/10.1007/978-3-662-44193-0_5⟩.

Evening, M., 2013. The Adobe Photoshop Lightroom 5 Book: The Complete Guide for Photographers. Pearson Education.

Fagan, J.C., 2010. Usability studies of faceted browsing: a literature review. Inf. Technol. Libr. 29 (2), 58.

Grey, J.M., 1977. Multidimensional perceptual scaling of musical timbres. J. Acoust. Soc. Am. 61 (5), 1270–1277.

Guest, S., Dessirier, J.M., Mehrabyan, A., McGlone, F., Essick, G., Gescheider, G., Fontana, A., Xiong, R., Ackerley, R., Blot, K., 2011. The development and validation of sensory and emotional scales of touch perception. Atten. Percept. Psychophys. 73 (2), 531–550.

Harrower, M., Brewer, C.A., 2013. Colorbrewer.org: an online tool for selecting colour schemes for maps. Cartogr. J.

Hearst, M., 2006. Design recommendations for hierarchical faceted search interfaces. In: Proceedings of the ACM SIGIR Workshop on Faceted Search, pp. 1–5.

Hearst, M.A., 2008. UIs for faceted navigation: recent advances and remaining open problems. In: Proceedings of the Second Workshop on Human-Computer Interaction and Information Retrieval (HCIR), pp. 13–17.

Hoggan, E., Brewster, S., 2007a. Designing audio and tactile crossmodal icons for mobile devices. In: Proceedings of the Ninth ACM International Conference on Multimodal Interfaces. pp. 162–169.
Hoggan, E., Brewster, S., 2007b. New parameters for tacton design. In: Proceedings of Extended Abstracts on Human Factors in Computing Systems (CHI EA), ACM, New York, NY, USA, pp. 2417–2422.
Hollins, M., Bensmaïa, S., Karlof, K., Young, F., 2000. Individual differences in perceptual space for tactile textures: evidence from multidimensional scaling. Percept. Psychophys. 62 (8), 1534–1544.
Hong, K., Lee, J., Choi, S., 2013. Demonstration-based vibrotactile pattern authoring. In: Proceedings of the Seventh International Conference on Tangible, Embedded and Embodied Interaction (TEI), pp. 219–222.
Israr, A., Zhao, S., Schwalje, K., Klatzky, R., Lehman, J., 2014. Feel effects: enriching storytelling with haptic feedback. ACM Trans. Appl. Percept. (TAP) 11, 11:1–11:17.
Jäschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., Stumme, G., 2007. Tag recommendations in folksonomies. In: Proceedings of European Conference on Principles of Data Mining and Knowledge Discovery, Springer, pp. 506–514.
Jason L. Huang, Paul G. Curran, J.K.E.M.P.R.P.D., 2012. Detecting and deterring insufficient effort responding to surveys. J. Bus. Psychol., 27 (1), pp. 99–114. URL: ⟨http://www.jstor.org/stable/41474909⟩.
Last.fm, 2016. URL: ⟨http://www.last.fm/music⟩. (Accessed 24 July 2016).
Levesque, V., Oram, L., MacLean, K.E., March 2012. Exploring the design space of programmable friction for scrolling interactions. In: Proceedings of IEEE Haptics Symposium (HAPTICS), pp. 23–30.
Lieberman, H., Paternò, F., Klann, M., Wulf, V., 2006. End-user development: an emerging paradigm. In: End User Development, Springer, pp. 1–8.
Lo, J., Johansson, R.S., et al., 1984. Regional differences and interindividual variability in sensitivity to vibration in the glabrous skin of the human hand. Brain Res. 301 (1), 65–72.
MacLean, K.E., 2008. Foundations of transparency in tactile information design. IEEE Trans. Haptics (ToH) 1 (2), 84–95.
MacLean, K., Enriquez, M., 2003. Perceptual design of haptic icons. In: Proceedings of EuroHaptics, pp. 351–363.
MacLean, K.E., Schneider, O., Seifi, H., 2017. Multisensory haptic interactions:understanding the sense and designing for it. In: The Handbook of Multimodal-Multisensor Interfaces, ACM Books.
Matlab, 2016. URL: ⟨http://thenounproject.com/⟩.
Nielsen, J., Molich, R., 1990. Heuristic evaluation of user interfaces. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 249–256.
NounProject, 2016. URL: ⟨http://thenounproject.com/⟩.
Obrist, M., Seah, S.A., Subramanian, S., 2013. Talking about tactile experiences. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1659–1668.
Okamoto, S., Nagano, H., Yamada, Y., 2013. Psychophysical dimensions of tactile perception of textures. IEEE Trans. Haptics (ToH) 6 (1), 81–93.
Pandora, 2016. URL: ⟨http://www.pandora.com/⟩. (Accessed 24 July 2016).
Peck, J., Childers, T.L., 2003. Individual differences in haptic information processing: the "need for touch" scale. J. Consum. Res. 30 (3), 430–442.
Ryu, J., Choi, S., 2008. posVibEditor: graphical authoring tool of vibrotactile patterns. In: Proceedings of IEEE International Workshop on Haptic Audio visual Environments and Games (HAVE), pp. 120–125.
Saul, G., Lau, M., Mitani, J., Igarashi, T., 2011. Sketchchair: an all-in-one chair design system for end users. In: Proceedings of the Fifth ACM International Conference on Tangible, Embedded, and Embodied Interaction (TEI), pp. 73–80.
Schneider, O.S., Israr, A., MacLean, K.E., 2015. Tactile animation by direct manipulation of grid displays. In: Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST), pp. 21–30.
Schneider, O.S., MacLean, K.E., 2014. Improvising design with a haptic instrument. In: Proceedings of IEEE Haptics Symposium (HAPTICS), pp. 327–332.
Schneider, O.S., MacLean, K.E., 2016. Studying design process and example use with macaron, a web-based vibrotactile effect editor. In: Proceedings of IEEE Haptics Symposium (HAPTICS), pp. 52–58.
Schneider, O.S., Seifi, H., Kashani, S., Chun, M., MacLean, K.E., 2016. Hapturk: crowdsourcing affective ratings of vibrotactile icons. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 3248–3260.
Seifi, H., Anthonypillai, C., MacLean, K.E., 2014. End-user customization of affective tactile messages: a qualitative examination of tool parameters. In: Proceedings of IEEE Haptics Symposium (HAPTICS), pp. 251–256.
Seifi, H., MacLean, K.E., 2013. A first look at individuals' affective ratings of vibrations. In: Proceedings of World Haptics Conference (WHC), pp. 605–610.
Seifi, H., Zhang, K., MacLean, K.E., 2015. Vibviz: organizing, visualizing and navigating vibration libraries. In: Proceedings of IEEE World Haptics Conference (WHC), pp. 254–259.
Smith, G., Czerwinski, M., Meyers, B., Robbins, D., Robertson, G., Tan, D.S., 2006. Facetmap: a scalable search and browse visualization. IEEE Trans. Vis. Comput. Graph. 12 (5), 797–804.
Swindells, C., Pietarinen, S., Viitanen, A., 2014. Medium fidelity rapid prototyping of vibrotactile haptic, audio and video effects. In: Proceedings of IEEE Haptics Symposium (HAPTICS), pp. 515–521.
TactileLabs, 2016. URL: ⟨http://tactilelabs.com/⟩.
Tam, D., MacLean, K.E., McGrenere, J., Kuchenbecker, K.J., 2013. The design and field observation of a haptic notification system for timing awareness during oral presentations. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), ACM, New York, NY, USA, pp. 1689–1698.
Ternes, D.R., 2007. Building Large Sets of Haptic Icons: Rhythm as a Design Parameter, and Between-Subjects MDS for Evaluation (Ph.D. thesis), The University of British Columbia.
Thompson, B., 2004. Exploratory and confirmatory factor analysis: understanding concepts and applications. American Psychological Association.
Turnbull, D., Barrington, L., Lanckriet, G.R., 2008. Five approaches to collecting tags for music. In: Proceedings of The International Society of Music Information Retrieval (ISMIR), vol. 8, pp. 225–230.
van Erp, J.B., Spapé, M.M., 2003. Distilling the underlying dimensions of tactile melodies. In: Proceedings of Eurohaptics, vol. 2003, pp. 111–120.
VibViz Dataset, 2016. URL: ⟨http://www.cs.ubc.ca/labs/spin/vibviz ⟩.
Yee, K.-P., Swearingen, K., Li, K., Hearst, M., 2003. Faceted metadata for image search and browsing. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 401–408.
Yohanan, S., Chan, M., Hopkins, J., Sun, H., MacLean, K., 2005. Hapticat: exploration of affective touch. In: Proceedings of the Seventh ACM International Conference on Multimodal Interfaces (ICMI), pp. 222–229.
Yohanan, S., MacLean, K.E., 2011. Design and assessment of the haptic creature's affect display. In: Proceedings of the Sixth ACM International Conference on Human-Robot Interaction (HRI), pp. 473–480.
Yong, A.G., Pearce, S., 2013. A beginner's guide to factor analysis: focusing on exploratory factor analysis. Tutor. Quant. Methods Psychol. 9 (2), 79–94.
Yoo, Y., Yoo, T., Kong, J., Choi, S., 2015. Emotional responses of tactile icons: effects of amplitude, frequency, duration, and envelope. In: Proceedings of IEEE World Haptics Conference (WHC), pp. 235–240.
Zhao, S., Schneider, O., Klatzky, R., Lehman, J., Israr, A., 2014. Feelcraft: crafting tactile experiences for media using a feel effect library. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST), ACM, Adjunct Publication, New York, NY, USA, pp. 51–52.
Zheng, Y., Su, E., Morrell, J.B., 2013. Design and evaluation of pactors for managing attention capture. In: Proceedings of IEEE World Haptics Conference (WHC), pp. 497–502.