Exploring User Experience with Image Schemas, Sentiments, and Semantics

Rossitza Setchi [ORCID], *Senior Member, IEEE,* and Obokhai K. Asikhia

Abstract—Although the concept of user experience includes two key aspects, experience of meaning (usability) and experience of emotion (affect), the empirical work that measures both the usability and affective aspects of user experience is currently limited. This is particularly important considering that affect could significantly influence a user's perception of usability. This paper uses image schemas to quantitatively and systematically evaluate both these aspects. It proposes a method for evaluating user experience that is based on using image schemas, sentiment analysis, and computational semantics. The aim is to link the sentiments expressed by users during their interactions with a product to the specific image schemas used in the designs. The method involves semantic and sentiment analysis of the verbal responses of the users to identify (i) task-related words linked to the task for which a certain image schema has been used and (ii) affect-related words associated with the image schema employed in the interaction. The main contribution is in linking image schemas with interaction and affect. The originality of the method is twofold. First, it uses a domain-specific ontology of image schemas specifically developed for the needs of this study. Second, it employs a novel ontology-based algorithm that extracts the image schemas employed by the user to complete a specific task and identifies and links the sentiments expressed by the user with the specific image schemas used in the task. The proposed method is evaluated using a case study involving 40 participants who completed a set task with two different products. The results show that the method successfully links the users' experiences to the specific image schemas employed to complete the task. This method facilitates significant improvements in product design practices and usability studies in particular, as it allows qualitative and quantitative evaluation of designs by identifying specific image schemas and product design features that have been positively or negatively received by the users. This allows user experience to be assessed in a systematic way, which leads to a better understanding of the value associated with particular design features.

Index Terms—Affect, computational semantics, image schema, interaction design, ontology, sentiment analysis, usability, user experience

✦

1 INTRODUCTION

USER experience is a multi-faceted concept that refers to the experiential and affective aspects of using a particular product, system, or service. Understanding user experience is critical to the design of high-quality interactive products because it helps identify the meaningful and valuable aspects of that interactive experience and improves the overall usability of these products.

User experience is defined as a person's perceptions and responses that result from the use or anticipated use of a product, system, or service [1]. The concept refers to all aspects of the interaction, including utility, ease of use, and efficiency, and is influenced by the context of use and individual beliefs, preferences, and emotions [2]. User experience is dynamic, situated, context-dependent, and subjective; it stems from a broad range of potential benefits users may derive from

a product [3]. It is measured through attributes related to the usability of a product (e.g., efficiency, perspicuity) and measures associated with perceptions (e.g., attractiveness, stimulation, dependability, and novelty). Studies suggest a clear relationship between need fulfilment and positive affect, with stimulation, relatedness, competence, and popularity being especially salient needs [4]. Therefore, the concept of user experience includes two key aspects: experience of meaning (usability) and experience of emotion (affect).

Usability is widely studied and well understood in the domains of Human-Computer Interaction (HCI) and design. It is defined as the extent to which a product can be used by a specific user to achieve a specified goal with effectiveness, efficiency, and satisfaction in a specified context of use [1]. Performance measures for usability evaluation include effectiveness, efficiency, guessability, flexibility, learnability, memorability, and satisfaction [1], [5], [6].

The second aspect, affect, is less well understood in the context of product design [2], [4], and the main focus is user satisfaction with products. Techniques used to assess user satisfaction include identifying subjective feelings through interviews and various physiological and motor measurements. However, there is a limited body of empirical work that explores in a holistic way both the usability and affective aspects of user experience. This is particularly important considering that affect could negatively influence a user's perception of usability.

- R. Setchi is with Cardiff School of Engineering, Cardiff University, 14-17 The Parade, Cardiff CF24 3AA, United Kingdom.
  E-mail: setchi@cardiff.ac.uk.
- O.K. Asikhia is with Edo State Institute of Technology and Management, Usen, P.M.B. 1104 Benin City, Edo State, Nigeria.
  E-mail: Asikhia.OK@esitmusen.edu.ng.

Recently, the use of image schemas has been proposed as a means for exploring intuitive interaction [7], [8]. Image schemas are defined as mental patterns that structure human experiences in the world [9]. However, most of the studies conducted using image schemas have predominantly focused on the usability aspect. This includes studies employing direct observation [10], [11], and think-aloud protocols [12], [13].

The *research hypothesis* of this paper is that users' attitudes and feelings can not only be identified but can be directly linked to the image schemas employed during their interactions with a product. The implication is the possibility of using image schemas to quantitatively measure user experience in its entirety, in terms of both usability and affect. This would lead to a better understanding of the value associated with particular design features.

To address this challenge, this paper proposes a *semantic method* for exploring users' experience that is based on using image schemas. The method includes an *ontology* of image schemas and a novel *algorithm* to identify the image schemas employed to complete a task and link a user's emotional response to the specific image schemas employed in the task. The method uses task-oriented and affective words employed by the users during their interactions with products.

The remainder of the paper is organized as follows. Section 2 provides background information and discusses relevant research in product design, including that dealing with user experience, image schemas, and sentiment analysis. Section 3 introduces the method developed, while Section 4 outlines its experimental evaluation. Section 5 consists of the results and the discussion. Finally, Section 6 offers conclusions and suggestions for future work.

2 BACKGROUND

This section discusses user experience, image schemas, and sentiment analysis in the context of product design.

2.1 User Experience

The concept of user experience encompasses both the functional (cognitive) and perceptual/emotional (affective) needs of the users [2]. It is about the relationship between the objective and the subjective, the internal and the external aspects constituting human–product interactions [14]. While the functional needs refer to the usability of the product, the affective aspect is concerned with the emotion that occurs as a result of the interaction with a product. The affective aspects of user experience are less well studied than the functional needs of users, as indicated by Hassenzahl and Tractinsky [2] who suggest that the emotional needs of the user must be better understood, defined, and operationalized.

This duality between reason and emotion has been studied by several prominent scholars, including Damasio [15] and Minsky [16]. For example, Damasio [15] argues that emotion is in the loop of reason; that is, emotion can assist in the reasoning process and on some occasions can even replace reason. He views emotion as delivering cognitive information, directly and via feelings. His theory suggests that when making decisions, people use physiological signals (or "somatic markers") associated with their past experiences. Several techniques have been developed to measure the affective state of the user; each technique has its strengths and weaknesses

[17]. The most comprehensive among them is the multi-component model for measuring affect, which accommodates five dimensions of affective measurement—cognitive appraisal, subjective feeling, physiological measurement, motor expression, and behavioral tendencies [18].

Forlizzi and Battarbee [19] classify the existing frameworks for structuring user experience as product-centered, user-centered, and interaction-centered models. The product-centered frameworks [20], [21] predominantly focus on the experience evoked by the product. They are mostly developed as guidelines to help designers incorporate aspects that can evoke pleasurable experiences when using the product. The user-centered frameworks [22], [23] focus on the aspects of the user's behavior that designers can embody into the design in order to facilitate the ease of use of the product. The interaction-centered frameworks [24], [25] focus on the experiences that the product evokes as a mediator between the designer and the user. Such frameworks incorporate aspects of the product-centered and the user-centered approaches.

Several researchers have identified the affective experience (emotion) as an integral component of the user experience [2], [15], [26]. In particular, Hekkert [26] describes user experience as the entire set of effects that are elicited by the interaction between a user and a product, including the degree to which all senses are gratified (aesthetic experience), the meaning attached to a product (experience of meaning or usability), and the emotion that is elicited (emotional experience). People naturally use their cognitive, motor, and affective skills to interact with objects [27]. The experience of meaning is linked to cognition; previous experience plays a key role in the association, interpretation, and retrieval of features from memory [26]. A product that is designed with the user's previous experience in mind does what the user expects, with minimal cognitive effort and within the shortest possible timeframe. Usability is measured using a number of criteria including effectiveness, efficiency, guessability, flexibility, learnability, memorability, and satisfaction [1], [5], [6]. The performance indicators commonly used are time on task for efficiency, accuracy for effectiveness, and preferences for measuring satisfaction.

Subjective feelings can be assessed through the use of questionnaires, rating scales, checklists, and analyses of verbal responses. Several tools have been developed for measuring the affective state based on subjective feelings. These include the Geneva Affect Label Code (GALC) [17], the Affect Grid [28], and dimensional emotion theory [29]. These tools are based on a fixed or free response of the user. For example, the affect grid, which is a fixed response tool, assesses the emotional state on a scale from 1 to 9 using such measures as pleasure–displeasure and arousal–sleepiness. Users are asked to rate their emotional state by selecting one field of the affect grid. However, the priming effect of the fixed response is one of the greatest limitations of the fixed response tool [17]. The free response is based on the assumption that people are aware to some extent of their emotional state and are able to describe their emotions [30]. GALC [17] is a free response tool that uses a limited number of emotional categories. It includes 36 categories of types of affect-related experiences and their associated word stems. For example, "disappointment" is a GALC category with 12 word stems: *comedown, disappoint*, *discontent*,

TABLE 1
Categories of Image Schemas [34]

Class	Category
Basic schemas	*Substance, object*
Space	*Up-down, left-right, near-far, front-back, center-periphery, contact, path, scale*
Containment	*Container, in-out, content, full-empty, surface*
Multiplicity	*Merging, collection, splitting, part-whole, count-mass, link, matching*
Process	*Iteration, cycle*
Force	*Diversion, counterforce, restraint removal, resistance, attraction, compulsion, blockage, balance, momentum, enablement*
Attribute	*Heavy-light, dark-bright, big-small, warm-cold, strong-weak, straight, smooth-rough*

disenchant*, *disgruntle**, *disillusion**, *frustrate**, *jilt**, *letdown*, *resign**, *sour**, and *thwart**.

2.2 Image Schemas

An image schema is defined as a dynamic pattern of organism-environment interaction that gives understanding to experience emanating from human bodily interaction with the physical world [9]. Image schemas are formed on the basis of interactions, pre-linguistic metaphors, and historical context.

For example, Johnson [9] describes the verticality image schema as providing "a basis for our up-down orientation based on different experiences such as perceiving a tree, our felt sense of standing upright, the activity of climbing stairs, forming a mental image of a flagpole, measuring our children's height, and experiencing water rising in a bathtub". These various experiences represent the abstract structure of the verticality schema. Over time, an association relating these recurring experiences with the observed relationship is established in the user's subconscious. These associations are "metaphoric extensions" of the recurring experience. Such experiences acquired by the user over a long period of time can form interesting patterns that can subsequently be used for interaction with minimum cognitive effort and in a quicker time frame [31]. For example, simple everyday experiences include seeing an object being inside or outside of a container, experiencing water rising in a river, or being at the center or periphery of a spatial scene [9], [32]. The words used to describe these experiences act as a gateway to the structure the brain has composed to store the corresponding information. In the above examples, *container*, *up-down*, and *center-periphery* are all image schemas developed from recurring and similar interactions with the world that have left traces in the brain.

Furthermore, each *container* image schema can be described as having an inside, outside, and a boundary that allows movement in and out of the container [33]. The *container* image schemas can be instantiated by the words *in, out, enter, emerge, come out, come in*, etc. Similarly, the *up-down* image schema is grounded in our experiences of gravity coded in words such as *up, down, top, increase, decrease, rise*, and *fall*, while *center-periphery* captures the experience of humans as being the perceptual center of their world [9]. Several image schemas, such as *up-down big-small*, *near-far*, and *left-right* have been studied extensively in cognitive

linguistics using verbal and nonverbal stimuli. These studies show that image schemas can be activated in audio, visual, and motor modalities.

The complete taxonomy of image schemas is still debated in the literature, but researchers agree that the image schemas can be grouped based on the relationships between cognitive domains. Table 1 includes an inventory of image schemas as categorized in [34] based on a large body of knowledge developed by Johnson [9], Lakoff [35], Clauster and Croft [36], and other researchers.

As mentioned in the introduction, image schemas have been used to explore the usability aspects of human interaction with products. Examples of products used in studies include cameras [37] and mobile phones [10]. One of the limitations of these studies is that the image schemas have been extracted manually from the experimental data. This is a significant limitation because the coding process is dependent on the expertise of the researcher conducting the study.

2.3 Sentiment Analysis in Product Design

Interacting with an interface is an experience that involves stimulus and response. Users' responses can be identified through their language and behavior. In the context of the think-aloud protocol widely used in product design evaluation experiments, humans use language to express their thoughts about their experience with a product. The opinions expressed by different people on the use of the same product vary based on many factors, one of which is their prior experience with the product. Several theories, including Slovic's "affect heuristic" theory [38] and Damasio's somatic marker theory [15], support the idea that user experiences are stored in the long-term memory along with emotion. During the retrieval process, the emotion acts as a reminder of the previously stored information. In the context of user interactions with products, emotional responses are generated based on the "affective tag" that has been attached to similar situations experienced in the past.

A paper aimed to clarify the terminological differences between affect, feelings, emotions, sentiments, and opinions [39] points out that 'sentiments involve emotional dispositions formed toward an object, and unlike emotions that are brief, sentiments about an object are enduring'. Sentiment analysis can be used to analyze the opinion expressed by the user and to identify their affective state during the interaction. It is defined as the use of natural language processing to characterize opinions, sentiments, evaluations, appraisals, attitudes, and emotions relating to entities such as products, services, organizations, individuals, issues, events, and other attributes in a source material [40].

Even though Natural Language Processing (NLP) in general has many unresolved problems related to deep language understanding, sentiment analysis is one of its success stories because it only requires understanding of the positive or negative sentiments of each sentence and the referent or topic in it [41]. It involves polarity classification, mood identification, emotion recognition, agreement detection, subjectivity analysis, and identifying the degree of positivity or negativity [40], [41].

Existing approaches to sentiment analysis include keyword spotting, lexical affinity, and statistical modelling [41]. As the name implies, keyword spotting classifies texts based

on the presence of affective words. The lexical affinity approach assigns an arbitrary word a probabilistic affinity for a particular emotion, while statistical modelling is based on machine learning applied to large data sources. All three approaches have their strengths and weaknesses. Although considered the most naïve approach, keyword spotting is popular due to its simplicity, accuracy, and low computational cost [41].

Sentiments can be discovered at document level, sentence level, or aspect level. In particular, aspect-level sentiment analysis aims to find sentiment-target pairs in a given text about an entity (i.e., the target) and its main characteristics (aspects) [42]. The four main approaches for identifying the sentiment-target pairs are machine learning, dictionary-based, statistical, and semantic [43].

A large amount of work has been undertaken to develop collections of affect-related words and reasoners that use such resources. Examples include the aforementioned GALC [17] and Affect Grid [28], as well as the Linguistic Annotation Scheme [44], WordNet-Affect [45] and SentiFul [47]. While earlier studies were dominated by statistical modelling and supervised learning, more recent work indicates a noticeable shift towards using concept-based approaches linked to common sense and linguistic knowledge bases. This trend has been also noticed by Schouten and Frasincar [42] who believe that the future of aspect-level sentiment analysis is semantically-rich concept-centric approaches.

In the domain of product design, sentiment analysis has been used in customer satisfaction studies [48] and marketing to analyze consumers' preferences for particular products and to establish a new product range [49]. It has been also employed in online personalized recommendation systems [50] and reputation management to analyze opinions about brands and products in product reviews and online fora [51], [52]. Most of the existing studies analyze the positive or negative opinions expressed by consumers in terms of overall opinion about products and particular usability functions. Such studies lack the in-depth analysis needed to link these opinions to specific design features.

To summarize, previous studies have shown the potential of image schemas in evaluating usability, but there is no work linking user experience with image schemas. In addition, sentiment analysis has been successfully applied in marketing and branding, but there is insufficient understanding of how it can be used to evaluate user experience. Moreover, none of the existing approaches to aspect-level sentiment analysis deal with metaphorical abstract categories (such as those used at the subconscious level of mental images), which are very challenging to express and therefore detect. This paper addresses these research gaps by proposing a method based on image schemas for assessing the affective aspects of users' interactions with products.

3 PROPOSED SEMANTIC METHOD FOR EXPLORING USERS' EXPERIENCE

3.1 Conceptual Model
The method supports evaluation studies of user experience based on the think-aloud protocol, which is a widely used tool in usability studies as it offers insight into non-observable aspects of interactions and can reveal misconceptions and

misinterpretations of design elements or instructions. Similar to direct observation, the think-aloud protocol allows the investigator to explore specific areas of interest in the context of a specified task. The utterances contain task-related and affect-related words and phrases. For example, the utterance *"The use of the down button to increase the time seems weird"* extracted from a participant's think-aloud transcript as part of this study contains the phrase *down button to increase*, which is related to a design feature and the task, while the word *"weird"* refers to the experience of using this particular design feature.

The identification of task-related and affect-related words requires the use of a different type of lexical resource. While a number of affect-related and generic lexical resources are currently available, the analysis of the task-related words and phrases requires the development of a new domain-specific ontology.

The proposed method aims at analyzing the utterances of the participants in the evaluation studies and linking them to particular image schemas used in the product designs (see Fig. 1 in which the novel contributions of this paper are indicated in italics). The method consists of a domain-specific *ontology of image schemas* and a novel *algorithm* developed to extract the image schemas used to complete a task; to determine, based on sentiment analysis, the affective words used and their polarity; and to link them to the specific image schemas used to complete the task. In addition to the ontology of image schemas, the method uses a general-purpose lexical ontology and a vocabulary of affective words.

3.2 Ontology of Image Schemas
Ontology is a branch of Philosophy that studies the nature and structure of reality. In Knowledge Engineering, however, the term refers to a domain model and is defined as a "formal, explicit specification of a shared conceptualization" [53], [54] and as "a logical theory designed to account for the intended meaning of a vocabulary" [55]. These two definitions support the rationale for developing a computational ontology of image schemas. In particular, in the context of this study, such an ontology would provide a simplified view and a structured representation of the knowledge about image schemas. Moreover, the ontology would provide a shared vocabulary of terms, which will allow reasoning about interactions with products and what the quality of particular product design features is

A number of methodologies is employed to develop and maintain domain ontologies. A major limitation of many of them is "tying the ontology closely to a specific task" [56]. This research uses METHONTOLOGY because it enables the construction of ontologies at the knowledge level (i.e., starting with conceptualization rather than implementation). Its standard development cycle includes the following steps: specification, knowledge acquisition, conceptualization, integration, implementation, evaluation, and documentation [57].

The specification phase in the development of the ontology of image schemas involves analyzing the existing classifications of image schemas and making decisions about the purpose, scope, formality, and granularity of the ontology to be developed. The starting point is the taxonomy shown in Table 1 [34], which includes seven classes and 41
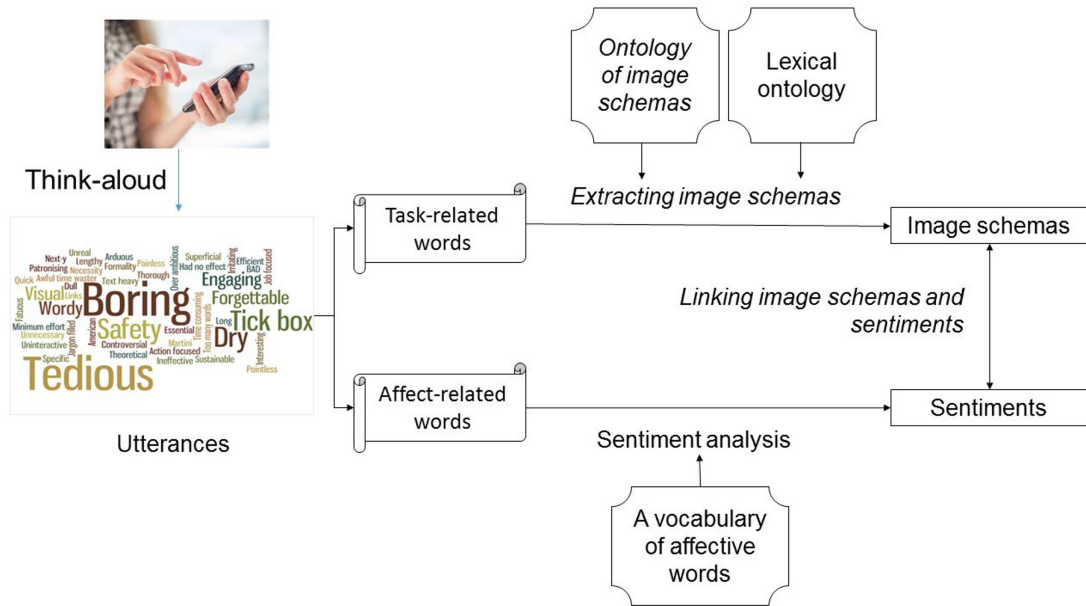
Fig. 1. Conceptual model of the proposed method. The elements in italics are the specific contributions of this paper.

categories. As the ontology developed in this research was conceived as an aid to identify image schemas that can be associated with user's experiences with physical products, it was decided to limit the scope to image schemas that are extensively used during interactions. These image schemas are: *containment, force, space, multiplicity,* and *attributes.* Table 2 shows the scope of the developed ontology, which includes five classes and twenty categories of image schemas. In terms of granularity, it was felt that the ontology needs *three* more levels in addition to classes and categories: a third level, which includes the most essential properties of each type of image schema; a forth level, which contains the linguistic concept associated with each property; and a fifth level, which contains the linguistic terms semantically related to each of those properties.

The knowledge acquisition phase involves using knowledge sources to identify the properties of each category of image schema. It has to be noted that not all metaphorical extensions are studied extensively or have precise definitions in the literature. It was found that one of the best resources available is the online catalogue of image schemas called *ISCAT* [58], which includes descriptions, definitions, references, evidence of experiential grounding, relationships to other image schemas, and a large amount of linguistic examples from the English and German languages. For example, *content* is listed in *ISCAT* as belonging to *containment* and is defined as "everything that is inside a *container*". Examples of metaphorical extensions of *content* are physical and emotional states, such as anger, fear, sadness, and happiness. Some among the many linguistic examples given for *content* are: "He has a pain in his shoulder" [31], "The smile left his face" [31], "She is boiling with anger" [59], and "My cold has gone from my head to my chest" [31]. *Content* is only related to *containment,* its parent class, but most image schemas are related to other entities. For example, *container* is related to *in-out, surface, blockage, center-periphery,* and *scale.*

The conceptualization phase includes using the definitions of the image schemas to associate properties with the image schema categories. A property in this context is a

lexical term that represents an essential element of an image schema category that could be related to the way a user interacts with a design feature. Table 3 contains examples of image schema categories, one from each class, and their corresponding properties. For example, *container* has four essential properties: interior, exterior, ingress, and egress; *path* is characterized by source, goal, and path (trajectory), etc. An important part of the conceptualization phase is linking properties to semantic concepts and linguistic terms.

The lexical resource used for this purpose is *OntoRo.* This is an electronic resource based on Roget's Thesaurus [60] and has been employed in several other studies related to design [61] and information retrieval [62], [63]. Moreover, the use of *OntoRo* is also justified by linguistic studies [36], which have found that many of the names of image schemas correspond to Roget's subcategories.

Similar to Roget's Thesaurus, *OntoRo* has a hierarchical structure comprising six classes, 39 sections, 95 subsections, and 990 semantic concepts. The linguistic terms within these concepts are organized in 'semicolon' groups that belong to one of the four parts of speech (i.e., noun, adjective, verb, and adverb). Within each part of speech, there are separate paragraphs that link related words semantically. For example, Roget's category *capacity* (found both under the heading *physical states* and the heading *qualities of matter*) parallels the content of a *container* image schema [36].

Next, the image schema properties are manually linked to the most suitable semantic concept in *OntoRo* (Table 4).

TABLE 2
Image Schemas Employed in the Ontology

Class	Category
Space	*Up-down, near-far, front-back, center-periphery, contact, path, scale, rotation*
Containment	*Container, in-out, content*
Multiplicity	*Part-whole, link, matching*
Force	*Attraction, compulsion, blockage, balance*
Attribute	*Dark-bright, big-small*

TABLE 3
Properties of Image Schemas

Class	Category	Definition [58]	Properties
Containment	*Container*	A *container* consists of a boundary that separates an interior (space enclosed by the boundary) from the exterior (the surrounding area/space not enclosed by the boundary) and, often, a portal (an opening in the boundary that allows movement between the interior and the exterior).	Interior, exterior, movement into (ingress), movement out (egress)
Space	*Path*	A *path* consists of a source or starting point, a goal or end-point, and a sequence of contiguous locations connecting the source with the goal.	Source, goal, path
Multiplicity	*Part- whole*	*Part-whole* is an image schema that consists of a whole, parts, and a configuration.	Whole, part, configuration
Attribute	*Big- Small*	*Big* is conspicuous in size, extent, position, or importance; it is significant. *Small* is insignificant, of a size that is less than normal or usual.	Big, small
Force	*Attraction*	*Attraction* is a force image schema in which a (passive) object exerts a force on another object, either physically or metaphorically, to pull it toward itself.	Attraction

Each of the 990 concepts in OntoRo is labelled with its number and the first word of all the words belonging to that concept (e.g., #224 interiority). All concept numbers in this study use a # tag. Most properties are polysemic words that can be linked to more than one concept. For example, the word *interior*, which is one of the properties of the image schema *container*, has six senses and can therefore be linked to six linguistic concepts: *#5: intrinsic, #58: component, #78: inclusion, #224 interiority, #344 land*, and *#553 painting*. However, only one of these concepts, *#224 interiority*, is related to *interior* as a space enclosed by a boundary, which is the description of *container*.

Even though concept disambiguation using algorithms developed by the authors in previous design studies [61] and information retrieval research [62], [63] is possible, this was considered unnecessary because this procedure is only performed as part of the conceptualization phase of the ontology.

Each concept in Roget's Thesaurus is linked to a large number of linguistic terms that are semantically related. For example, concept *#224 interiority* is semantically linked to 191 words and phrases grouped in five paragraphs—two paragraphs with nouns, one with adjectives, one with verbs, and one with adverbs. It must be noted that not all words and phrases make immediate sense in the domain of product design and interactions, but Roget's Thesaurus is a general-purpose lexical resource and some associations are only applicable within a certain context. However, image schemas are used as metaphorical extensions where physical and emotional states are entities within a person (see the previously given example "My cold has *gone from my head to*

TABLE 4
Image Schemas' Categories and Related *OntoRo* Concepts

Class	Category	Concept
Space	*Up-down*	#209: height; #210: lowness; #308: ascent; #309: descent
	Near-far	#199: distance; #200: nearness
	Front-back	#237: front; #238: rear
	Center-periphery	#52: whole; #225: centrality; #234: edge
	Contact	#45: union
	Path	#68: beginning; #69: end; #624: way
	Scale	#27: degree
	Rotation	#315: rotation
Containment	*Container*	#224: interiority; #223: exteriority; #297: ingress; #298: egress
	In-out	#224: interiority; #223: exteriority; #297: ingress; #298: egress
	Content	#224: interiority
Multiplicity	*Part-whole*	#52: whole; #53: part; #58: component; #243: form
	Link	#47: bond
	Matching	#18: similarity; #462: comparison
Force	*Attraction*	#291: attraction
	Compulsion	#740: compulsion
	Blockage	#702: hindrance
	Balance	#28: equality
Attribute	*Dark-bright*	#417: light; #418: darkness; #425: color
	Big-small	#32: greatness; #33: smallness; #195: size; #196: littleness; #638: importance; #639: unimportance

Image Schema Ontology

| Class | Space | Containment | Multiplicity | Force | Attribute |

| Category | | Container | In-Out | Content | |

| Property | | Interior | Exterior | Ingress | Egress |

| Linguistic Concept | | #224: interiority | #223 | #297 | #298 |

Linguistic Terms

(191 words and phrases)
n,1: interiority, interior, inside, indoors, inner surface, under surface, endoderm, skin, sapwood, heartwood, wood, inmost being, heart's blood, soul, heart, centre, breast, bosom, centrality, inland, Midlands, heartland, hinterland, up-country, the nitty gritty, pith, marrow, substance, subsoil, substratum, base, permeation, pervasion, presence, interjacency, interspace, interval, deepness, cave, pit, pothole, penetralia, recesses, innermost recesses, depth, endogamy, marriage, introversion, intrinsicality, self-absorption, egoism, egotism, egocentrism, egocentricity, egomania, selfishness, introvert, egoist, egotist, inmate, dweller, internee, prisoner;
n,2: insides, contents, inner man, inner woman, interior man, internal organs, viscera, vitals, heart, ticker, lungs, lights, liver, kidneys, spleen, offal, meat, bowels, entrails, innards, guts, pluck, tripe, intestines, colon, und um, back passage, abdomen, belly, paunch, underbelly, womb, uterus, stomach, tummy, maw, chest, breast, bosom, solar plexus, gland, endocrine, cell, organism;
a, 1: interior, internal, inward, intrinsic, inside, inner, innermost, midmost, central, inland, up-country, removed, domestic, home, vernacular, intimate, familiar, known, indoor, intramural, shut in, enclosed, inboard, built-in, inwrought, endemic, residing, deep-seated, deep-rooted, ingrown, fixed, intestinal, visceral, intravenous, subcutaneous, interstitial, interjacent, inward looking, introvert, intrinsic, endo-, endogamous, endogenous;
v, 1: be inside, be internal, be within, lie within, lie beneath, be at the bottom of, show through, be visible, hold within, hold, comprise, place within, embed, insert, keep inside, intern, imprison, enfold, embay, enclose, internalize, absorb;
adv, 1: inside, within, in, deep in, deep down, inly, inwardly, intimately, deeply, profoundly, at heart, within side, within doors, indoors, at home, en famille, chez, at the sign of.
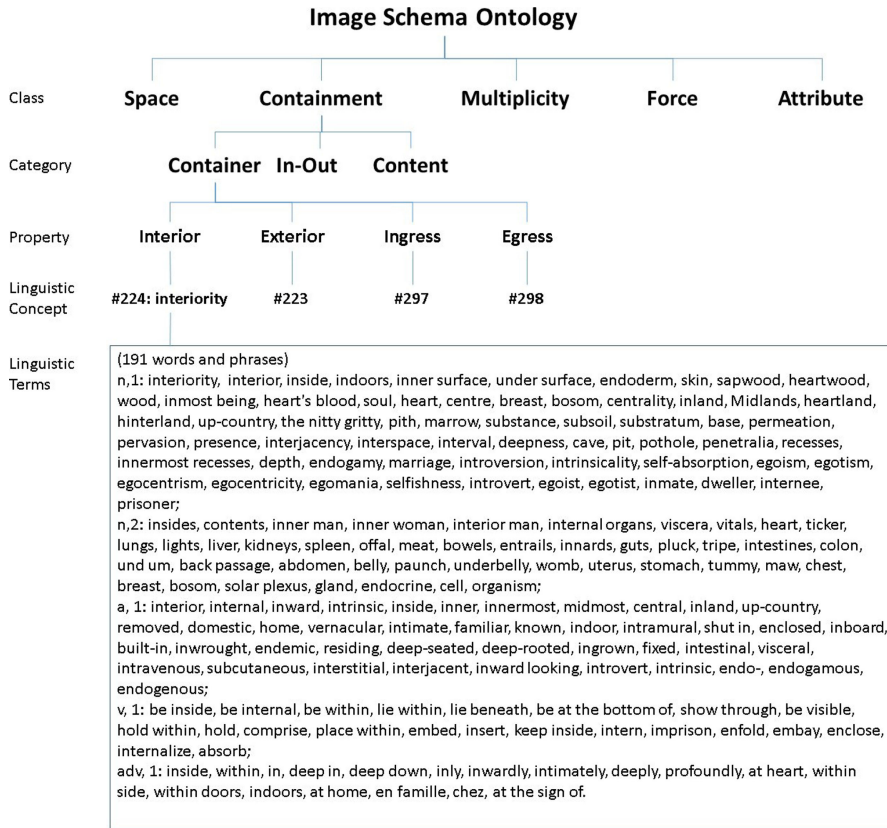
Fig. 2. Ontology of image schemas.

my chest" [31]), so it is appropriate that concept #224 *interiority* includes words related to internal organs, such as chest, belly, lungs, and guts.

Fig. 2 shows the structure of the developed ontology. It contains five classes, 20 categories, and 46 properties linked to the same number of linguistic concepts. Each concept is semantically related to between 120 and 450 words and phrases. The ontology also captures the semantic similarity between some image schemas. For example, *center-periphery* is semantically linked, according to the Catalogue of Image Schemas [58], to *container*, *in-out*, *near-far*, and *scale*. Table 3 shows the image schemas implemented in the ontology and their linguistic concepts. The number of related concepts corresponds to the number of properties associated with each image schema.

3.3 Algorithm for Extracting Image Schemas and the Sentiments Associated with Them

Fig. 3 shows the algorithm, which starts with transcribing the utterances of the participants in the think-aloud experiment. An utterance is an uninterrupted chain of spoken (or written) language. The transcription can be performed semi-automatically or automatically, where the utterances are differentiated when a longer pause occurs. The prosody (e.g., intonation, rhythmic quality, stress patterns, etc.) that accompanies speech is not included in the language transcript. All utterances are numbered, and these numbers are used later to link the extracted image schemas with the associated sentiment.

The transcripts are then pre-processed, which involves removing all words with very high frequency in the English

language (e.g., *the, of, to, a, this, haveç know, look*, etc.), as well as interjections, that is, words that have no grammatical meaning but are used in spoken language to signify emotions (e.g., *ahh, ahhh, aww, nah*, etc.). In addition, the words in plural forms (e.g., *buttons*), gerund forms (e.g., *looking*),
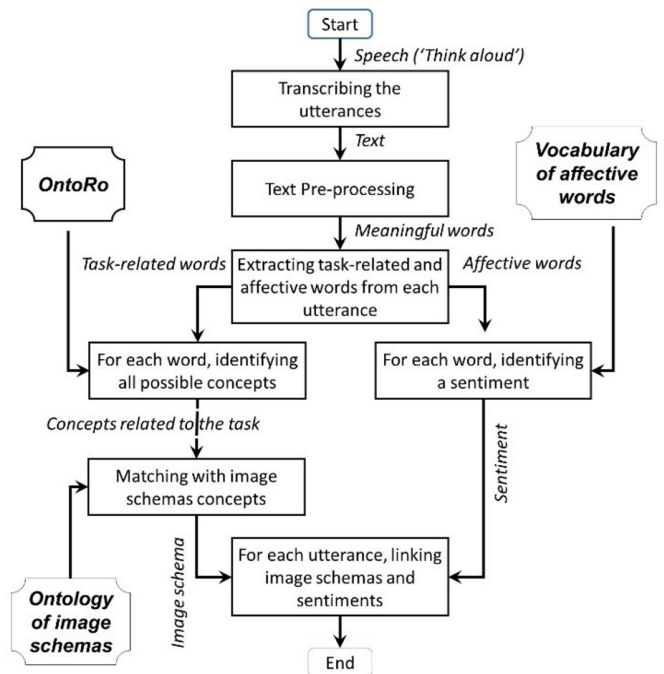
Fig. 3. Algorithm for extracting image schemas and the sentiments associated with them.

TABLE 5
Semantic Expansion

Word	No. of Concepts	Related Concepts
Press	16	*74: crowd; 198: make smaller; 216: flatter; 218: be supported, 258: smooth; 279: impel; 322: weigh; 587: press; 599: resolute; 612: incite; 678: acting; 691: advice; 740: compel; 761: request; 786: take away; 889: caress.*
Part	13	*46: disunite; 53: part; 55: incompleteness; 58: component; 263: open; 294: diverge; 296: depart; 410: melody; 412: vocal music; 589: reading matter; 594: acting; 622: function; 783: portion.*
Change	15	*15: differ; 19: make unlike; 131: middle age; 142: fitfulness; 143: change; 147: be turned to; 150: quid pro quo; 152: changeableness; 152: vary; 172: unproductiveness, 178: influence; 228: wear; 229: doff; 797: coinage; 797: money.*
Mode	6	*7: state, 86: numeration, 410: melody (concord), 610: habit, 624: way, 848: fashion (etiquette).*

and past tense (e.g., *asked*) are stemmed. However, rather than using the standard Porter stemmer [64], which removes all suffixes, this work uses the semantic stemmer SETS [63], which has shown good results with *OntoRo*. SETS keeps the words in the form they are used (e.g., *connecting* is not stemmed to *connect*) as the inflected words may have different meaning and different semantic representation (i.e., can be linked to different *OntoRo* concepts).The Porter stemmer creates more polysemic stems, which makes concept disambiguation—especially when working with short sentences—extremely difficult and inaccurate. Every word that occurs in an utterance is searched first in *OntoRo* using its morphological form as it is used by the participant in the experiment. If the word is found in *OntoRo*, it is considered stemmed, and the SETS algorithm proceeds to the next word. If the word does not occur in it, the algorithm proceeds with the first, second, third, etc., step of the Porter stemmer, repeating the search in *OntoRo* after each step [63]. Finally, if the algorithm does not find a stem, the word is considered a named entity and is not linked to any concepts. The SETS algorithm carefully avoids treating phrases as words (e.g., *present time* is treated as a phrase and not two individual words).

The next step is extracting the task-related and affective words in each stemmed utterance. This can be accomplished semi-automatically or automatically. A simple practical approach is removing all affective words using a collection of affect-related words and treating the rest as task-related words.

For example, the utterance *»1: I have to press this part to change the mode, nice* is given a number (1) and a tag (»), which are used in this research to code utterances. The utterance is preprocessed; the words *I, have, to, this, to, the* are ignored; and the stemmed utterance is recorded as *»»1: press part change mode nice (*the tag *»»* indicates that the utterance has been stemmed). There is one affective word in this sequence, so two more strings are recorded: *»»1.t: press part change mode* (for task-related words) and *»»1.a: nice* (for affective words).

The next step is identifying all possible senses for each task-related word. The step uses a semantic algorithm [61], [63], which identifies all possible senses and their corresponding concept numbers in *OntoRo*. Table 5 shows this process of *semantic expansion* using the example discussed above. The next step is matching the task-related words with the image schemas by comparing the lists of concept numbers related to each word using the image schema ontology (Table 3). In the previously used example, each of the 16 semantically related concepts associated with the task word *press* is matched with the image schema ontology

(Table 3, Fig. 2). Only one of the 16 concepts is related to an image schema, and this is concept #*740: compulsion*, related to the image schema *compulsion* of class *force*. Similarly, *part* is associated with #*53: part*, image schema *part-whole* of class *multiplicity*, and *mode* is related to #*624: way*, image schema *patç* class *space*. No direct match is found for the task-related word *change*. This step is recorded by generating a string *»»1.t: 740 53 0 624*.

Next, the affective words are analyzed in terms of expressing positive or negative sentiments using a vocabulary of affective words. Then a new string is recorded, which converts the string representing the affective words in an utterance (*»»1.a: nice*) into a string representing the sentiments expressed (*»»»1.a: positive*).

The final step involves linking the two strings and associating the identified image schemas with the sentiment expressed (*»»»1.t: 740 53 0 624) → (»»»1.a: positive*).

The algorithm is implemented in C++.

4 EXPERIMENTAL VALIDATION OF THE PROPOSED METHOD

The proposed method is evaluated using an empirical study involving 40 participants completing a set task with two different alarm clocks. Both of these products had design flaws, so the purpose of the experiment was to verify if the method is sensitive to identifying the specific issues experienced by the users during their interactions. In total, 80 observations were recorded.

4.1 Participants

Forty participants were recruited for the study. The breakdown shows 13 females, with age ranging from 22 to 43 years (mean = 30.54, SD = 6.05), and 27 males, with age ranging from 22 to 45 years (mean = 30.31, SD = 5.07). The group is multicultural, but this aspect was not considered important in the study. Fig. 4 shows a participant interacting with *Product 2*.

4.2 Apparatus

Two multifunctional alarm clocks were used in the study. These products were chosen because they provided many interaction opportunities based on a wide range of image schemas. The first alarm clock had seven features—changing the mode, setting time, setting timer, enabling sound, up and down buttons, and display. The second alarm clock also had seven features—switching on the alarm, setting up alarm, setting up hour, minute, timer, snooze button (which

Fig. 4. A participant during the experimental study.

stops the alarm and sets it to ring again a short time later), and display. These features are common to most digital products and people feel confident in interacting with them.

The equipment used in the experiment included the following:

- (i)   Four digital cameras. These cameras recorded participants' interactions from four angles.
- (ii)   Noldus Observer XT software for capturing, analyzing, and presenting behavioral data [65].

4.3 Procedure

The participants in the study were given the task to set the alarm clock to 11:30 am. Before the start of the experiment, all clocks were set to 3:00 am and were in alarm mode. The participants were not told how to complete the tasks and were not given any product documentation or instruction manuals.

The correct execution of the tasks was extracted by the researchers from the product documentation. It included a sequence of three subtasks for *Product 1* (activate the time mode, set the hour, set the minutes) and five subtasks for *Product 2* (activate the alarm mode, activate the set option, set the hour mark, set the minute mark, and activate the set time).

Fig. 5 shows the design of *Product 1*. It consists of a display at the front and five buttons (alarm, time, hour, minutes, and snooze) located at the top of the product. In addition, an alarm switch is located at the right side of the product. As indicated in the instruction manual for the alarm clock, the task requires the execution of three subtasks:

- (1).   Activate the time mode. This requires pressing the alarm switch up to enter the time mode.
- (2).   Set the hour. This step involves holding the time button down and pressing the hour button until it gets to 11 am.
- (3).   Set the minutes. This involves holding the time button down and pressing the minute button until it gets to 30.

The product description for the same alarm clock adds more detail. It states that "the time mode environment is displayed on the top left corner of the screen with a dot sign."

Fig. 5. Design features of *Product 1*.

The text describing the subtasks and the product was used to identify the image schemas that have to be activated according to the designer's intent.

Based on the information in the instruction manual, the image schemas involved in completing subtask 1 are *compulsion*, *path*, and *container*. The *compulsion* image schema is associated with the action word "press". The external force in the form of the users physically pushing the button causes the passive entity (alarm clock) to move from the alarm mode to the time mode. The *path* image schema is extracted from the movement along a path initiated by the force involving pressing. The alarm mode corresponds to the start point; the action *press* initiates a movement along a path to the end point, which is the normal mode. The *container* image schema is extracted from the word *enter*. The normal mode represents the location out of the container; the action word *press* initiates a movement through the portal into the container, which represents the alarm mode. Two additional image schemas were extracted from the product description. These are *part-whole* and *matching*. The central location of the screen creates a spatial relationship with other features. The screen represents the *part of the whole* configuration. The dot symbol was used to represent the time mode, and the requirement to find the dot symbol expects the use of the *matching* image schema. Therefore, five image schemas were expected to be used for the completion of subtask 1: *compulsion*, *path*, *container*, *part-whole*, and *matching*.

The participants were timed and observed during the task; all interactions were video-recorded. The participants were encouraged to verbalize their thoughts during their interaction with the products. The participants' verbal responses were recorded and later transcribed. The time to complete the task and the number of errors were processed using the Observer XT10 software. A structured questionnaire was completed by the participants immediately after completion of the task. The participants were also asked to rate each product in terms of their overall satisfaction using a five-point Likert scale. The study lasted approximately 30 minutes per participant.

5   RESULTS AND DISCUSSION

In total, 80 interactions were recorded as a result of 40 participants completing tasks with two products. The participants needed between 85 and 502 seconds to complete the task with *Product 1* and between 34 and 396 seconds to complete the task with *Product 2*. The number of words captured

TABLE 6
A User's Interactions with *Product 1 (Subtask 1)*

Subtask	Direct observation	Think-aloud protocol
Activate the time mode	The participant was observed to search for the switch controlling the time and alarm mode. The participant pushed (indicating the use of *compulsion*) the switch forward (use of *path*). The participant was observed to examine the content of the screen (*container*) to see if there is an indication of the time mode (*matching*).	*"I have to look carefully to know where to begin. The arrangement of the features is lovely. Okay. I have seen the switch. I have to press the switch forward to move into time mode. I can't really see anything that represents time mode. This is unexpected."*
Set the hour	The participant was observed to try out different buttons and different combinations, occasionally looking at the screen to see if there was a change.	*"I have pressed almost all the buttons and nothing seems to be changing. This is very difficult. Aha, okay, I got it. You will need to hold the alarm and press the hour to set the hour to 11."*
Set the minutes	The participant was observed to search for the time and the minute buttons. The participant repeated the same procedure as in the previous subtask.	*"The minutes are quite simple to set. Hold the time and press the minutes until it gets to 30."*

through the think-aloud protocol were between 30 and 185 for *Product 1* and between 27 and 162 words for *Product 2*.

Two aspects of the subtasks for *Product 2* were the most problematic for the majority of the participants: increasing the hour setting and accessing the alarm mode. When attempting to increase the hour, the majority of the participants were affected by the implementation of the up-down image schema. The *down* feature was used to increase the time while an *up* feature to decrease it. Participants' mental model of experience associating an increase with *up* and decrease to *down* was violated and though the subtask was completed quickly, many errors were also made. The second aspects of the difficulties experienced by the majority of the participants in *Product 2* involve accessing the alarm mode. While trying to enter the alarm mode, the majority of the participants used variations of the statement "I am looking for how to get into alarm mode", which speaks of the difficulty in locating the right button for the sub-task. Participants had problems differentiating between the buttons that were similar in labels, which increased their cognitive load of finding the right way to complete the task.

Even though *Product 1* had fewer subtasks, the majority of the participant spent more time trying to figure out how to complete the task. For example, to get into alarm mode, participants had to combine the alarm and hour buttons, which appeared very confusing to the majority of the participants during the experiment.

Overall, the majority of the participants in the study made several mistakes while completing the task using the two products, but the majority of the participants were quicker to recover from the mistakes after several trials for product 2 than for *Product 1*.

Table 6 shows the type of data recorded for one of the participants in the study as a result of direct observation and the think-aloud protocol.

The analysis of the direct observation data in Table 6 shows that the participant attempted to use *compulsion*, *path*, *container*, and *part-whole* to activate the time (sub-task 1). However, the image schema *matching* has manifested itself as a problem.

The analysis of the think-aloud data for the same subtask shows that the meaningful task-related words (i.e., those linked to image schemas) used are: *arrangement* (related to concept #53: *part* and image schema *part-whole*), *press* (#740: *compulsion* and image schema *compulsion*), *move in* (#297: *ingress* and image schema *container*), *mode* (# 624: *way* and image schema *path*), and *represent* (#18: *similarity* and image schema *matching*). The affective words accompanying these task-related words are *lovely* (*positive* sentiment associated with *pleasure/enjoyment*, linked in this example to *part-whole*) and *weird* (*negative* sentiment indicating *surprise*, related in this context to the image schema *matching*). For this participant, distinguishing between the time mode and the alarm mode, which involves using a *matching* image schema, appears to have been a problem (the affect word *unexpected* is used to describe the interaction). In contrast, the arrangement of the features that are linked to the *part-whole* image schema is referred to positively as *lovely*. No affective words were used in relation to the image schemas *compulsion*, *path*, and *container*.

In total, 65 affective words were used in the whole experimental study, including 32 positive words from nine GALC categories of positive sentiments and 33 negative words from six GALC categories. Table 7 lists the affective words used by the participants, the category they belong to, and their valence (P-positive, N-negative).

As previously mentioned, the image schemas expected to be used in subtask 1 according to the designer's intent, as extracted from the manual and the product description, were *compulsion*, *path*, *container*, *part-whole*, and *matching*. All these image schemas were identified by the two methods of direct observation and the think-aloud protocol. However, the image schema *matching* was identified as problematic.

Table 8 shows the data analysis for the experiment. It includes data from direct observations and the think-aloud protocol for all 40 participants interacting with the two products. The success rate for each participant is calculated as the ratio of the correctly used image schemas over the total number of image schemas. The total number is a sum of the correctly used, the incorrectly used (attempted but causing difficulties), and the unidentified (not attempted). The success rate of all participants was 37.81% with *Product 1* and 25 percent with *Product 2* (as mentioned previously, both products had design flaws and so a low success rate was expected). These low percentages indicate that the method is very sensitive, even when applied to a relatively simple type of product that most people attempt to operate without reading a manual.

As mentioned earlier, the participants were asked to indicate their level of satisfaction with the two products as

TABLE 7
Affective Words Used by the Participants

Affective Word	Affect Category (GALC)	Valence
Fascinated, wondering,	Admiration	P
Funny	Amusement	P
Comfortable, satisfying	Contentment	P
Disappointed, frustration	Disappointment	N
Disliked	Disgust	N
Clear, easy, feel, instinctive, intuitive, like, perceived, simple, useful, user, friendly	Feeling	P
Confident	Hope	P
Interest	Interest	P
Awful, distracted, unpleasant	Irritation	N
Attractive, wish	Longing	P
Bad, challenge, incompatible, poor, uncertain, unfamiliar, useless,	Negative	N
Classical, compatible, expected, explicit, familiar, fine, good, helpful, obvious, preferable	Positive	P
Appealing, beautiful, lovely, nice	Pleasurable/ Enjoyment	P
Amaze, strange, surprising, unexpected, weird	Surprise	N
Awkward, complex, complicated, confusing, difficult, disagree, discomfort, hard, obscure, problematic, terrible, tough, tricky, uncomfortable	Tension/ stress	N

TABLE 8
Users' Interactions with Products 1 and 2 (All Subtasks)

Participant	Product 1				Product 2			
	Correctly used image schemas	Incorrectly used image schemas	Unidentified image schemas	Success rate	Correctly used image schemas	Incorrectly used image schemas	Unidentified image schemas	Success rate
1	5	1	2	0.625	2	-	6	0.500
2	2	3	3	0.250	2	1	5	0.500
3	5	-	3	0.625	1	2	5	0.125
4	4	1	3	0.500	1	2	5	0.125
5	5	-	3	0.500	2	2	4	0.500
6	3	2	5	0.375	5	-	3	0.625
7	4	2	2	0.500	2	-	6	0.250
8	2	2	4	0.250	2	1	5	0.250
9	4	-	4	0.500	2	-	6	0.250
10	4	2	2	0.500	2	2	4	0.250
11	2	2	4	0.250	1	1	6	0.125
12	3	2	3	0.374	2	-	6	0.250
13	2	1	5	0.250	2	-	6	0.250
14	1	2	5	0.125	1	-	7	0.125
15	2	-	6	0.250	3	-	5	0.375
16	3	1	4	0.375	2	-	6	0.500
17	3	-	5	0.375	1	-	7	0.125
18	2	2	4	0.250	1	-	7	0.125
19	3	-	5	0.375	1	-	7	0.125
20	4	1	3	0.500	2	-	6	0.250
21	3	3	5	0.375	2	-	6	0.250
22	2	1	5	0.500	2	-	6	0.250
23	3	-	5	0.375	2	-	6	0.250
24	1	-	7	0.125	1	-	7	0.125
25	3	1	4	0.375	1	-	7	0.125
26	4	-	4	0.500	1	-	7	0.125
27	5	-	3	0.625	1	1	6	0.125
28	3	2	3	0.375	1	-	7	0.125
29	3	1	4	0.375	1	1	6	0.125
30	1	2	5	0.125	1	-	7	0.125
31	2	2	4	0.250	-	-	-	-
32	2	1	5	0.250	1	-	7	0.125
33	4	-	4	0.500	3	-	5	0.375
34	5	1	2	0.625	1	1	6	0.125
35	2	1	5	0.250	3	-	5	0.375
36	2	1	5	0.250	1	-	7	0.125
37	2	2	4	0.250	2	-	6	0.500
38	2	1	5	0.250	1	-	7	0.125
39	3	2	5	0.375	2	-	6	0.250
40	5	2	1	0.625	2	-	6	0.250

part of the structured questionnaire they completed after the experiment. The levels of satisfaction were used to group the participants in three groups: low (group 1), medium (group 2), and high satisfaction (group 3).

Next, a statistical analysis was conducted to determine if there were significant differences in the number of positive sentiments expressed by the three groups. Only participants who completed the tasks successfully were considered in this analysis. Out of all forty participants in the study, twenty seven participants successfully completed the task with Product 1, and thirty participants successfully completed the task with Product 2.

A Shapiro-Wilks test ($p = 0.257$) showed that the sentiments expressed in relation to Product 1 were approximately normally distributed in the three groups with effect size ($d = 0.500$). Homogeneity of variance was violated, as assessed by the Levene test of equality of variance ($p = 0.040$), and so separate variances and the Welch correction were used. There were statistical differences in the three groups: $F(2, 24) = 6.381$ ($p = 0.060$). A Tukey post hoc test revealed there was a statistically significant difference between the sentiments expressed in the high and low satisfaction groups ($p = 0.0060$) and the high and medium groups ($p = 0.036$). No statistically significant difference was found between the low and medium satisfaction groups ($p = 0.733$).

Similarly for Product 2, the Shapiro-Wilks test ($p = 0.135$) showed that the sentiments expressed were normally distributed in the three groups, with effect size $d = 0.416$ and a Levene test of equality of variance $p = 0.862$. There were statistically significant differences in the three groups: $F(2, 27) = 3.619$ ($p = 0.001$). A Tukey post hoc test revealed a statistically significant difference between the sentiments expressed in the high and low satisfaction groups ($p = 0.037$). No statistically significant difference was found between the high and medium groups ($p = 0.168$) and the low and medium groups ($p = 0.739$). These results confirmed that participants with high satisfaction ranking were more likely to use positive affective words to describe the image schemas employed in the
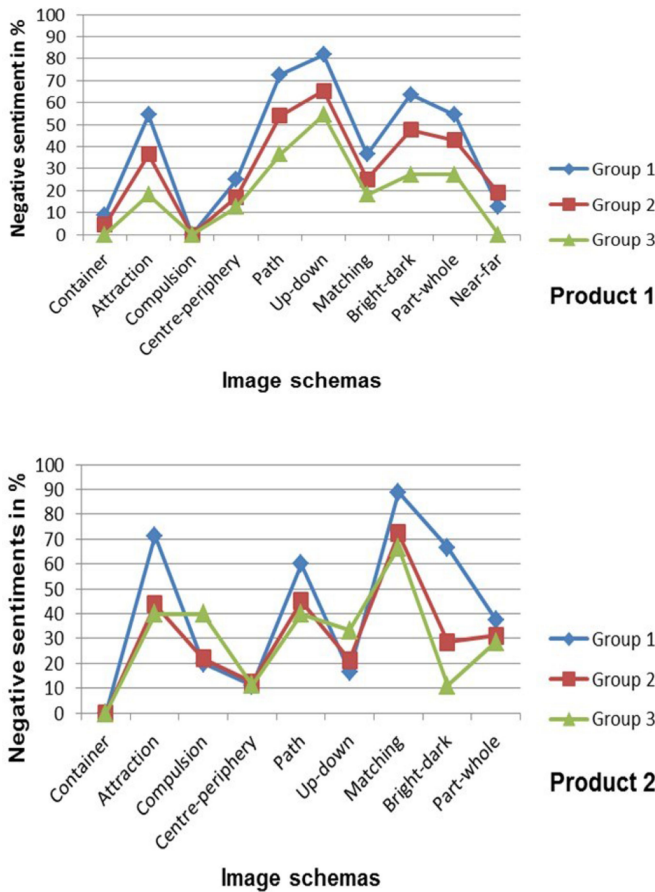
Fig. 6. Negative sentiments expressed during interactions with *Product 1 and Product 2.* The participants are grouped according to the level of their satisfaction: low (group 1), medium (group 2), and high satisfaction (group 3).

interactions. As expected, there is a positive correlation between the level of satisfaction and the number of positive sentiments expressed through affective words.

The most significant result, however, is demonstrated in Fig. 6, which shows the negative sentiments expressed towards individual image schemas. The image schemas associated with more negative sentiments were *path, bright-dark, attraction,* and *matching* for *Product 1* and *path, up-down,* and *bright-dark* for *Product 2.* These image schemas were also identified in the questionnaire completed by the participants as causing difficulties and as responsible for the relatively low levels of satisfaction with the two products.

A closer examination of the image schemas responsible for the task difficulties with *Product 1* reveals that the *path* image schemas affected the performance of the participants. The designer's idea of combining buttons to set the alarm time (subtask 1) created interaction problems for many participants. Similarly, the majority of the participants applied trial and error in subtask 2, which resulted in more errors and longer task completion time. The *matching* image schema affected the participants in terms of the appearance of the icon used to represent the alarm mode. The use of a "dot" symbol to represent the alarm mode on the screen was not meaningful to the users.

For Product 2, the *up-down* image schema affected the participants in terms of the spatial arrangement of the features. The participants made more errors and had increased

search times as a result of placing the feature to increase the time in the reverse location. Furthermore, the *path* image schema (comprising the feature used to get into the "alarm" mode) was again problematic. The majority of the participants affirmed that they were used to the features "set" and "mode" to assess menus in most digital products. This particular product comes with mode and set features.

Participants were split between the use of the two buttons to get into "alarm" mode. Many errors and longer times were observed in this subtask (subtask 1). The *bright-dark* and *matching* image schemas affected the participants in terms of the appearance of the features (color, label, and the icon used). The poor contrast in terms of the color used for the label and the background created visual perception problems for the majority of the participants. In addition, the icon used to represent the "alarm" mode was presented with a poor color choice that did not attract the attention of the majority of the participants.

These results suggest that the method used in this study can be used for in-depth analyses of user interfaces and their design elements. The method allows identifying the effects of using particular image schemas, and if the effect is not the one intended by the designer, changes can be recommended to improve the product design. Furthermore, the mental model of the participants regarding the functionality of the system based on the sequencing of action and system response can be evaluated based on this approach. This can help in detecting interaction difficulties and potential design flaws in the early phase of design.

6 CONCLUSIONS AND FUTURE WORK

This is the first study that experimentally investigates the link between image schemas and affect. This is significant because sentiment analysis in product design has been primarily used in customer satisfaction studies, marketing, branding, and reputation management. This study proposes a method that can be used to analyze interactions and the quality of specific design features.

The main contribution of the paper is the semantic method, which identifies and links the sentiments expressed by the user relating to the specific image schemas used in a task. Another important innovative aspect of the method is the domain-specific ontology of image schemas developed for this study. In addition to this ontology, the method uses general-purpose and specialized linguistic resources.

The empirical study conducted to validate the approach shows that the user experience can be directly linked to the image schemas used during an interaction with a product. This can potentially lead to significant improvement in design for intuitive use, as the approach directly links experience to the specific image schemas employed in the design. The ontology of image schemas can be used in any study investigating intuitive interactions in the area of design for intuitive use.

Some of the limitations of this study suggest opportunities for further research. First, the ontology of image schemas can be extended with more categories and more relationships between semantic concepts and image schema properties. Second, other, more advanced, sentiment

analysis techniques and more advanced linguistic resources and reasoners can be employed. Finally, the semantic and affective analysis can be extended to include interjections, that is, words that have no grammatical meaning, but are used in spoken language to signify emotions and the prosody that accompanies speech.

REFERENCES

[1] ISO 9241: *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)*. Part 11: Guidance on Usability, International Organization for Standardization, 1998.

[2] M. Hassenzahl and N. Tractinsky, "User experience—A research agenda," *Behaviour Inf. Technol.*, vol. 25, no. 2, pp. 91–97, 2006.

[3] E. L.-C. Law, V. Roto, M. Hassenzahl, A. P. O. S. Vermeeren, and J. Kort, "Understanding, scoping and defining user eXperience: A survey approach," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2009, pp. 719–728.

[4] M. Hassenzahl, S. Diefenbach, and A. Göritz, "Needs, affect, and interactive products – facets of user experience," *Interacting Comput.*, vol. 22, no. 5, pp. 353–362, 2010.

[5] J. Nielsen, *Usability Engineering*. Rochester, NY, USA: AP Professional, 1993.

[6] P.W. Jordan, *Introduction to Usability*. Park Drive, U.K.: Taylor and Francis, 1998.

[7] J. Hurtienne, J.H. Israel, and K. Weber, "Cooking up real world business application combining physicality, digitality, and image schema," in *Proc. Int. Conf. Tangible Embedded Interaction*, 2008, pp. 239–246.

[8] O. K. Asikhia, R. Setchi, Y. Hicks, and W. Andrews, "Conceptual framework for evaluating intuitive interactions based on image schemas," *Interacting Comput.*, vol. 27, no. 6, pp. 287–310, 2015, doi:10.1093/iwc/iwu050.

[9] M. Johnson, *The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason*. Chicago, IL, USA: University of Chicago Press, 1987.

[10] A. Britton, R. Setchi, and A. Marsh, "Intuitive interaction with multifunctional mobile interfaces," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 25, pp. 87–196, 2013.

[11] A. N, Antle, G. Corness, and A. Bevans, "Springboard: Designing image schema based embodied interaction for an abstract domain," *Human-Computer Interaction Series: Whole Body Interaction*, D. England, ed. Berlin, Germany: Springer, 2011, pp. 7–18.

[12] P. P. Maglio and J. Matlock, "The conceptual structure of information space," in *Social Navigation of Information Space*, A. Munro, D. Benyon, and K. Hook, eds. London, U.K.: Springer, 1999, pp. 155–173..

[13] J. Hurtienne and P. Langdon, "Keeping warm in winter: Image-schematic metaphors and their role in design of central heating controls," in *Proc. 4th Int. Conf. German Cognitive Linguistics Assoc.*, 2010, pp. 53–54.

[14] F. Picillo and G. Cascini, "A framework for user experience, needs and affordances," *Des. Studies*, vol. 35, pp. 160–179, 2014.

[15] A.R. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain*. New York, NY, USA: Putnam, 1994.

[16] M. Minsky, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York, NY, USA: Simon Schuster, 2006.

[17] K. R. Scherer, "What are emotions? and how can they be measured?" *Soc. Sci. Inf.*, vol. 44, no. 4, pp. 695–729, 2005.

[18] K. R. Scherer, "On the nature and function of emotion: A component process approach," *Approaches to Emotion*, K. R. Scherer and P. Ekman, eds. New York, NY, USA: Erlbaum, 1984, pp. 293–317.

[19] J. Forlizzi, and K. Battarbee, "Understanding experience in interactive systems," in *Proc. Conf. Designing Interactive Syst.*, 2004, Art. no.. 261.

[20] L. Alben, "Quality of experience: Defining the criteria for effective interaction," *Interactions*, vol. 3, no. 3, pp. 11–15, 1996.

[21] V. Jääskö and T. Mattelmäki, "Observing and probing," in *Proc. Int. Conf. Designing Pleasurable Products interfaces*, 2003, pp. 126–131.

[22] A. Mäkelä and J. Fulton Suri, "Supporting users' creativity: Design to induce pleasurable experiences," in *Proc. Conf. Affective Human Factors*, 2001, pp. 387–391.

[23] M. Hassenzahl, "The thing and I: Understanding the Relationship between User and Product," *Funology: From Usability to Enjoyment*, M. Blythe, C. Overbeeke, A. F. Monk, and P. C. Wright, eds. Norwell, MA, USA: Kluwer, 2003, pp. 31–42.

[24] J. Forlizzi and S. Ford, "The building blocks of experience: An early framework for interaction designers," in *Proc. Des. Interactive Syst.*, 2000, pp. 419–423.

[25] A. Mäkelä and T. Mattelmäki, "Collecting stories on user experiences to inspire design - A pilot," in *Pleasure with Product*, W. S. Green and P. W. Jordan, eds., Park Drive, U.K.: Taylor Francis, 2002, pp. 333–344.

[26] P. Hekkert, "Design aesthetics: Principles of pleasure in product design," *Psychology Sci.*, vol. 48, no. 2, pp. 57–172, 2006.

[27] S. A. G. Wensveen, *A Tangibility Approach to Affective Interaction*. Ultimo, NSW, Australia: Univ. Technol., 2005.

[28] J. A. Russell, A. Weiss, and G. A. Mendelssohn, "The affect grid: A single-item scale of pleasure and arousal," *J. Personality Soc. Psychology*, vol. 57, pp. 493–502, 1989.

[29] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.

[30] A. Mehrabian, "Framework for a comprehensive description and measurement of emotional states," *J. Genetic Psychology*, vol. 121, no. 3, pp. 339–361, 1995.

[31] G. Lakoff and M. Johnson, *Metaphor We Live By*. Chicago, IL, USA: Univ. Chicago Press, 1980.

[32] J. Hurtienne, K. Klöckner, S. Diefenbach, C. Nass, and A. Maier, "Designing with image schemas: Resolving the tension between innovation, inclusion and intuitive use," *Interacting Comput.*, vol. 27, no. 6, pp. 235–255, 2015.

[33] B. Hampe and J. E. Grady, eds., *From Perception to Meaning: Image Schema in Cognitive Linguistics*. Berlin, Germany: Mouton de Gruyter, 2005, pp. 1–12.

[34] J. Hurtienne and L. Blessing, "Design for intuitive use-testing image schema theory for user interface design," in *Proc. 16th Int. Conf. Eng. Des.*, 2007, Art. no. 386.

[35] G. Lakoff, *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago, IL, USA: Univ. Chicago Press, 1987.

[36] T. C. Clausner and W. Croft, "Domains and image schemas," *Cognitive Linguistics*, vol. 10, pp. 1–32, 1999.

[37] W. Kuhn and A. U. Frank, "A formalisation of metaphors and image schema in user interfaces," in *Cognitive and Linguistic Aspect of Geographical Space*, D. M. Mark and A. U. Frank, eds. Dordrecht, The Netherlands: Kluwer Academic, 1991, pp. 419–434.

[38] P. Slovic, M. Finucane, E. Peters, and D. G. MacGregor, "The affect heuristic," in *Heuristics and Biases: The Psychology of Intuitive Judgment*, T. Gilovich, D. Griffin, and D. Kahneman, eds. Cambridge, U.K.: Cambridge Univ. Press, 2002, pp. 397–420.

[39] M. Munezero, C. S. Montero, E. Sutinen, and J. Pajunen, "Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text," *IEEE Trans. Affective Comput.*, vol. 5, no. 2, pp. 101–111, Apr.-Jun. 2014.

[40] B. Liu, *Semantic Analysis and Opinion Mining*. Williston, VT, USA: Morgan Claypool Publisher, 2012.

[41] E. Cambria, B. Schuller, B. Liu, H. Wang, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 15–21, Mar.-Apr. 2013, doi:10.1109/MIS.2013.30.

[42] K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 3, pp. 813–830, Mar. 2016.

[43] M. Tsytsarau and T. Palpanas, "Survey on mining subjective data on the web," *Data Mining Knowl. Discovery*, vol. 24, no. 3, pp. 478–514, 2012.

[44] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language Resources Evalu.*, vol. 39, no. 2, pp. 165–210, 2005.

[45] C. Strapparava and A. Valitutti, "WordNet-Affect: An affective extension of WordNet," in *Proc. 4th Int. Conf. Language Resources Evalu.*, 2004, pp. 1083–1086.

[46] E. Cambria, D. Olsher, and D. Rajagopal, "SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis," in *Proc. 28th AAAI Conf. Artif. Intell.*, 2014, pp. 1515–1521.

[47] A. Neviarouskaya, H. Prendinger, and M. Ishizuka, "SentiFul: A lexicon for sentiment analysis," *IEEE Trans. Affective Comput.*, vol. 2, no. 1, pp. 22–36, Jan.-Mar. 2011.

[48] J. Zhu, H. Wang, M. Zhu, B. K. Tsou, and M. Ma, "Aspect-based opinion polling from customer reviews," *IEEE Trans. Affective Comput.*, vol. 2, no. 1, pp. 37–49, Jan.-Mar. 2011.

[49] S. N. Kolhe, S. V. Darade, P. B. Suryavanshi, P. S. Ahire, and H. D. Sonawane, "User feedback analysis system using natural language processing and artificial intelligence," *Int. J. Recent Innovation Trends Comput. Commun.*, vol. 3, no. 9, pp. 5574–577, 2015.

[50] A. C. M. Fong, B. Zhou, S. Hui, J. Tang, and G. Hong, "Generation of personalized ontology based on consumer emotion and behavior analysis," *IEEE Trans. Affective Comput.*, vol. 3, no. 2, pp. 152–164, Apr.-Jun. 2012.

[51] A. Abbasi, H. Chen, and A. Salem, "Sentiment analysis in multiple languages: feature selection for opinion classification in web forums," *ACM Trans. Inf. Syst.*, vol. 26, no. 3, pp. 12:1–12:34, 2008.

[52] B. Fong and J. Westerink, "Affective computing in consumer electronics," *IEEE Trans. Affective Comput.*, vol. 3, no. 2, pp. 129–131, Apr.-Jun. 2012.

[53] R. Studer, R. Benjamins, and D. Fensel, "Knowledge Engineering: Principles and Methods," *Data Knowl. Eng.*, vol. 25, no. 1–2, pp. 161–198, 1998.

[54] T. R. Gruber. "A translation approach to portable ontologies," *Knowl. Acquisition*, vol. 5. no. 2, pp. 199–220, 1993.

[55] N. Guarino, D. Oberle, and S. Staab, "What is an Ontology?" *Handbook on Ontologies*. Berlin, Germany: Springer, 2009.

[56] D. Jones, T. Bench-Capon, and P. Visser, "Methodologies for ontology development," in *Proc. 15th FIP World Comput. Congr.*, 1998, pp. 62–75.

[57] M. Fernández-López, A. Gómez-Pérez, and N. Juristo, "METHONTOLOGY: From ontological art towards ontological engineering," in *Proc. Spring Symp. Ontological Eng. AAAI*, 1997, pp. 33–40.

[58] ISCAT. (2007). [Online]. Available: http://iscat.stefciu.de

[59] A. Stefanowitsch, "Words and their metaphors: A corpus-based approach," in *Corpus-Based Approaches to Metaphor and Metonymy*, A. Stefanowitsch and S. T. Gries, eds. Berlin, Germany: Mouton de Gruyter, 2006, pp. 61–105.

[60] A. Kennedy and S. Szpakowicz, "Evaluation of automatic updates of Roget's thesaurus," *J. Language Model.*, vol. 2, no. 1, pp. 1–49, 2014.

[61] R. Setchi, Q. Tang, and I. Stankov, "Semantic-based information retrieval in support of concept design," *Adv. Eng. Inf.*, vol. 25, no. 2, pp. 131–146, 2011.

[62] S. A. Fadzli and R. Setchi, "Concept-based indexing of annotated images using Semantic DNA," *Eng. Appl. Artif. Intell.*, vol. 25, no. 8, pp. 1644–1655, 2012.

[63] I. Stankov, D. Todorov, and R. Setchi, "Enhanced cross-domain document clustering with a semantically enhanced text stemmer (SETS)," *Int. J. Knowl.-based Intell. Eng. Syst.*, vol. 17, no. 2, pp. 113–126, 2013.

[64] M. F. Porter, "An algorithm for suffix stripping," *Program*, vol. 14, no. 3, pp. 130–137, 1980.

[65] Noldus Observer XT 10, Innovative Solutions for Behavioral Research, 2010, pp. 5–39.

Rossitza Setchi (SM'14) received the MEng degree in mechanical and electrical engineering from Moscow Technological University, Moscow, Russia and the PhD degree from Cardiff University, Cardiff, United Kingdom. She leads the Mechanics, Materials, and Advanced Manufacturing Research Theme at Cardiff University and the Knowledge Engineering Systems (KES) Research Group. She received the 1999 Literati Club Award for Excellence, the 2004 IMechE Donald Julius Groen Prize, and the 2015 KES Award for outstanding contribution. She is an associate editor of the *International Journal of Systems Science* and the *International Journal of Knowledge and Systems Science*. She is a chartered engineer, a chartered IT professional, fellow of the Institution of Mechanical Engineers, fellow of the Institution of Engineering and Technology, and fellow of the British Computer Society. She is a senior member of the IEEE, and a member of several technical committees including IFAC TC 4.5 on Human-Machine Systems and IFIP TC 5.4 on Computer-Aided Innovation. Her current research interests include intelligent systems, robotics, artificial intelligence, and knowledge engineering.

Obokhai K. Asikhia received the MEng degree in industrial engineering from the University of Benin, Nigeria and the PhD degree from Cardiff University, Cardiff, UK. Currently, he is senior lecturer in the Department of Mechanical Engineering, Edo State Institute of Technology and Management, Nigeria. He is a corporate member of the Nigerian Society of Engineers (NSE) and registered member of Council for the Regulation of Engineering Practice in Nigeria (COREN). His research interests include design for intuitive use and semantic systems.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.