Title Page

Paper Title: Investigating users' reactions to surprising products

Author Names and Affiliations:

Author 1 Name and Surname: Niccolò Becattini
Author 1 ORCID: 0000-0002-1641-3796
Author 1 - Affiliation: Politecnico di Milano – Dipartimento di Meccanica, Via Privata Giuseppe La Masa, 1 – 20156 Milano, Italy.

Author 2 Name and Surname: Yuri Borgianni
Author 2 ORCID: 0000-0002-5284-4673
Author 2 - Affiliation: Free University of Bozen-Bolzano, Faculty of Science and Technology, Piazza Università, 5 - 39100 Bolzano, Italy.

Author 3 Name and Surname: Gaetano Cascini
Author 3 ORCID: 0000-0003-1827-6454
Author 3 - Affiliation: Politecnico di Milano – Dipartimento di Meccanica, Via Privata Giuseppe La Masa, 1 – 20156 Milano (Italy).

Author 4 Name and Surname: Federico Rotini
Author 4 ORCID: 0000-0002-1676-0835
Author 4 - Affiliation: Università degli Studi di Firenze – Dipartimento di Ingegneria Industriale, Via di Santa Marta, 3 – 50139 Firenze (Italy).

Corresponding Author:

Niccolò Becattini
Address: Politecnico di Milano – Dipartimento di Meccanica, Via Privata Giuseppe La Masa, 1 – 20156 Milano, Italy
Phone: +39 02 2399 8680
E-mail: niccolo.becattini@polimi.it

Abstract: The study of surprising product features is crucial for designing products that potentially trigger attention and curiosity. Through a tailored questionnaire, this study gathered reactions from 100 respondents to solutions which were considered to be surprising. The data about surprise emergence and its modalities were processed using a situated FBS-based cognitive framework, shifted to the perspective of the user/observer. Data analysis shows that FBS variables and the related cognitive processes are suitable for describing similarities and differences in the reasoning path of users when surprise emerges. This confirms that individually pre-conceived expectations are crucial to surprise emergence and that these expectations relate to functional, behavioural or structural variables with similar mechanisms that depend on thinking processes triggered by product features.

Keywords:

1. User behaviour

2. Creativity
3. Evaluation
4. Design cognition
5. Design Research

Investigating users’ reactions to surprising products

Abstract

The study of surprising product features is crucial for designing products that potentially trigger attention and curiosity. Through a tailored questionnaire, this study gathered reactions from 100 respondents to solutions which were considered to be surprising. The data about surprise emergence and its modalities were processed using a situated FBS-based cognitive framework, shifted to the perspective of the user/observer. Data analysis shows that FBS variables and the related cognitive processes are suitable for describing similarities and differences in the reasoning path of users when surprise emerges. This confirms that individually pre-conceived expectations are crucial to surprise emergence and that these expectations relate to functional, behavioural or structural variables with similar mechanisms that depend on thinking processes triggered by product features.

Design is devoting increasing attention to surprise as a potential driver of product success (Gross, 2016). Despite the dual nature of surprise (positive and negative) (Desmet, 2012), surprising products are generally considered intriguing and raise interest whatever their actual compliance with people's goals (Reisenzein et al., 2019). In this respect, Desmet (2003) discusses the potential in designing products that elicit paradoxical emotions, i.e. those that mix positive and negative aspects. More specifically, surprising artefacts can stand out in a competitive environment in which small performance increments are hardly captured by customers (e.g. Borgianni and Hatcher, 2017). The leveraging of surprise in design enables product experiences which result in attachment and a low likelihood of rapid disposal (Grimaldi, 2017). Moreover, surprise can draw people's attention and thus, achieve market success; in this respect, Hutter and Hoffmann (2014) discuss the unexpectedness of advertising ambient media. In addition, the presence of surprising elements in design is supposed to be beneficial for the design process itself. While the introduction of surprising information enhances the quality of design processes by fostering inspiration, enriching idea generation, and encouraging problem reframing (Stompff, 2016; Dankfort et al., 2018; Hatcher et al., 2018), the aim to design surprising products results in a significant boost of individually perceived creativity (Gotzsch, 2018).

Despite this evidence, surprise has not actually found its place in design research. While it is agreed that surprise is a result of the violation of expectations (see Section 1), the design field still lacks a proven and shared taxonomy of cognitive triggers to achieve the infringement of people's expectations. As better documented in the next section, existing studies have not yet translated what a user perceives into rules or guidelines designers can use to generate ideas for surprising products. Surprise emergence and its effects have been widely investigated in several domains, especially psychology, which has led to established knowledge and models. However, those cannot be replicated by designers when designing, as product features triggering surprise have not been systematically studied hitherto. A precondition to replicate surprise emergence through designed products is the definition and consolidation of a cognitive model describing the interpretation of product distinctive variables triggering surprise, thus in a designerly way and through constructs and concepts designers commonly master.

The present paper proposes the means for characterising users' reactions; through comments on potentially surprising products and catalogues, these outputs are framed according to acknowledged design schemas, such as the variables of the Function-Behaviour-Structure (FBS) ontology. The underlying assumption, to be verified here through a bespoke protocol, is that the FBS ontology, being representative of any kind of design variables (e.g. parameters designers can change), is also suitable for comprehensively describing product features (what the user identifies, interprets and perceives) which trigger surprise.

1. Background

The background includes four subsections. The first subsection presents an overview of the concept of surprise in emotion psychology and, markedly, in appraisal theory. The second subsection documents how appraisal theory has been valued within design with a specific focus on surprise. The third subsection outlines how design research has treated surprise so far. The fourth subsection explains the inadequacy of the most diffused approaches to extract information relevant to the design of surprising products.

1.1 Emotion psychology and surprise within appraisal theory

Appraisal theory's core is the determination of emotions based on multiple checks on the effects of the stimuli that have engendered emotions themselves (Scherer, 2001). In order to inform the scope of any research, these appraisal criteria or checks have to be as few as possible and, at the same time, have to distinguish the main families of emotional states. To date, there is no standard sample of appraisal variables used by scholars, as documented in Moors et al. (2013). For instance, in Scherer (2001), the number of checks is four: relevance, implications, coping potential and normative significance, which are supposed to take place sequentially, although this happens in a very short time. Other appraisal theories include novelty or variations thereof, such as suddenness and unexpectedness (Moors et al., 2013). In any case, it has been asserted that stimuli and events that surpass a given relevance threshold or deserve attention because of their novelty should always be processed further. Indeed, an emotion episode is circumscribed by the time at which "the synchronization or coherence of organic subsystems" are found beyond a certain threshold (Scherer, 2019). More specifically, emotion episodes are not limited to appraisal, which elicits the category of emotion experienced by an individual, but include subsequent actions (Scherer and Moors, 2019). Those are constituted by actions addressed to cope with the event that has triggered the emotion along with corresponding physiological responses, the integration of the new information that the event has engendered and, in some cases, the communication of the emotion episode.

Surprise is seldom present within the most diffused emotion episodes, but it is showcased in several studies within appraisal theory. This is in line with the relevance attributed to newly processed information and unexpectedness. In Rosenman et al. (1996), unexpectedness clearly emerges as the best appraisal by which to distinguish surprise from other emotions, whereas novelty, unfamiliarity and uncertainty are not good predictors of surprise. The same study confirms the double nature of surprise in terms of pleasantness, since both motive-consistent and -inconsistent events can give rise to surprise, classifiable as positive and negative, respectively. This partially conflicts with previous studies; for instance, Ellsworth and Smith (1988) link surprise to uncertainty beyond the relevance of agency, i.e. surprising events are mostly dictated by external circumstances. In contrast, the hedonic neutrality of surprise, with reference to pleasantness, is confirmed in Ellsworth and Smith (1988). Although researchers are not in complete agreement about this conclusion in the literature (Reisenzein et al., 2019), the most common stance foresees the existence of positive and negative surprise (see also Desmet, 2012). In Reisenzein et al. (2019), the link between surprise and the appraisal of unexpectedness is even more apparent. On the one hand, other possible appraisals and evaluations, such as novelty and valence, are to be excluded as elements capable of characterising surprise. On the other hand, it is highlighted that the appraisal of unexpectedness represents the most reliable method for the determination of a surprising event; this rules out, among others, the possibility of extrapolating the occurrence of surprise based on physiological or motoric responses. Overall, in Reisenzein et al. (2019), the emergence of surprise can mainly be linked to events that are discrepant with schemata formed on beliefs or social rules. This violation is capable of shifting attention towards the causal event and interrupting ongoing processes, in compliance with Scherer's (2019) definition of emotions provided above.

Therefore, surprise seems to possess the unique capability of orienting people's attention despite the possible lack of congruence or incongruence with their goals. It can be inferred that the occurrence of surprise can attract people's attention to products, although their main functions do not match a clear need or a predetermined desire for use and possession.

1.2 Appraisal theory in design and its implications for the elicitation of surprise by products

Appraisal theory is thought to be the most comprehensive framework for explaining the mechanisms relevant to emotional design, i.e. the processes that target the design of products intended to elicit specific emotions (Demir et al., 2009; Ho and Siu, 2012). However, given that appraisal theory is rooted in emotion psychology, its constructs are not easily transferrable to design, in particular because of its failure to focus on relationships with products (Demir et al., 2009). Desmet (2003) has developed a model which is presented along with relevant examples, in which the interplay between people's concerns, the product and an appraisal process leads to emotions. Markussen (2009) remarks on how this process is not static and is limited to the first exposition of products, and that accumulated experience with products and technical artefacts may lead to mixed emotions. However, the model is descriptive (Chowdhury et al., 2015) and the underlying mechanisms cannot be easily replicated during the design of a product. Indeed, in line with Demir et al. (2009) and with reference to design, it is challenging to identify the appraisal components involved in human-product interaction relevant to arouse emotion. This implies that research on emotional design has gradually focused on more specific emotions, while early studies had a tendency to target undifferentiated positive emotions (Desmet, 2012). When it comes to creating a taxonomy of emotions relevant to products, Desmet (2012) includes surprise, which is said to be experienced in response to sudden events and in violation of expectations and beliefs; this confirms the findings from studies on appraisal theory. In particular, in Desmet (2012), infringed expectations of surprising products include "what products are, what they do (how they respond during interaction)", beyond their performances.

From this perspective, the unexpected in products is plainly the determinant to be investigated in order to understand surprising reactions. Moreover, the level of unexpectedness (of the same product features) should be able to surpass a threshold that implies the shifting of attention and surprise reactions. Since, as stated above and openly in Desmet (2003), surprise does not necessarily relate to people's motives and individual goals (not known a priori), the design of surprising products can be based on targeting what might overall be considered unexpected.

1.3 Surprise in the design field

Beyond the positive or negative emotional response due to surprise (which is not the focus of this paper), it is worth defining design tools oriented to product features which trigger surprise. Those tools should be better based on cognitive models which are specifically oriented towards design, allowing designers to devise those features that violate expectations and fruitfully adopt similar strategies leveraging design variables to generate new surprising products. The starting point is the analysis of previous contributions discussing surprise in the design field. The main interest for surprise in design lies in the evaluation and assessment of design creativity (Becattini et al., 2017a) – here, the debated relationship between novelty and surprise mirrors arguments within appraisal theory. However, for the scope of the present paper, it is particularly relevant to analyse previous design contributions whose main thrust is the characterisation of surprising products and the strategies that facilitate unexpectedness. In this sense, the contribution authored by Ludden et al. (2008) can be considered a seminal work. The scholars illustrate design strategies that generate inconsistencies or incongruities embedded in products, in particular by means of visual illusions. This trajectory for causing unexpectedness is stressed in a subsequent publication (Ludden et al., 2012), in which articulated experiments that leverage visual-tactile inconsistencies are carried out. Rodríguez Ramírez (2014) investigates the strategies followed by outstanding industrial designers to devise potentially surprising products. The work shows a set of possible strategies that extend the previous sample of known techniques focused on visual illusions. Potential strategies include the use of shapes, components and physical principles that are commonly attributed to different worlds or industries. Unexpected modifications of structures and behaviours also contribute to the emergence of surprise according to Becattini (2017a). This empirical study also sheds light on phenomena that do not directly relate to technical choices, such as design intentions to violate human expectations in terms of habits, ethics and aesthetics. In a certain sense, the dimensions involved in potentially surprising products cover the whole FBS ontology, which will be used in the discussion below as a reference to investigate unexpected features and processes triggered by surprising products.

However, none of the previously mentioned studies has provided evidence of exhaustiveness, independence and repeatability of the features that allow products to be surprising. Even more remarkably, all these baseline studies introduce categories of surprising features that originate from scholars' understanding, but which are not based on users' explanations of surprise emergence. Additionally, in Becattini et al. (2017a), in which independent subjects are involved, participants were asked whether considered features could have caused surprise or could be considered pertinent to specific products, but in a retrospective way.

All these limitations justify the objective to explore how the perception of unexpectedness arises when caused by surprising products through the analysis of users' concurrent explanations.

1.4 Reasons to fine-tune a bespoke online questionnaire in the present study

Several approaches are presented in design literature to extrapolate information concerning design processes and/or deliverables. A thorough overview of this topic goes beyond the scope of the present paper; readers can refer to Ahmed (2007) and Dinar et al. (2015). The main approaches relevant to the present research are summarised below.

- *Protocol analyses* (Gero and Mc Neill, 1998) involve in-depth analysis of designers' behaviour and benefit from acknowledged schemas to characterise their activities, gestures, intentions, etc. Their main drawback is the significant effort required to process protocol data (Ahmed, 2007). As a consequence, they typically involve small samples of designers to ensure that the analysis of the protocols is sustainable (Hay et al., 2016).
- *Interviews* target the extraction of information from experts and their design practices, as in Rodríguez Ramírez (2014), Crilly (2015) and O'Hare et al. (2018). Here, the extracted information is difficult to classify and analyse, although interviews are

conducted in a (semi-) structured way. Indeed, retrospective effects potentially give rise to misrepresentations of design acts or practices conducted in the past (Wilson, 2013).

- *Ethnographic studies* in the design field date back to the 1990s (Bucciarelli, 1998), but they are currently facing new challenges due to their application in increasingly complex organisations (Petersen and Buch, 2016). Their key advantage is the observation of designers in their "natural" environment while performing their everyday job instead of in lab conditions. The observation of a design experience that is maximally unbiased by external factors can also be carried out when potential users undergo investigations (Mattelmäki et al., 2014), which makes it possible to extract people's experiences, practices and emotions relevant to product development.
- When the perspective shifts from designers to users, especially when considerable numbers of subjects are required, the so-called *Voice-of-the-Customer* (VoC) is commonly used. However, due to the market orientation of the VoC, these approaches are mostly focused on mapping customers' preferences, needs and expectations and, as such, they are only indirectly related to the design variables (Becattini, 2017b).

As the objective of the paper is to characterise surprised reactions to products (designed by someone else), the number of subjects (i.e. designers or users taking part in the above experimental approaches) is a critical factor. This rules out the adoption of the above first three approaches, as they are suitable for small samples only due to practical reasons. Conversely, the classical VoC techniques are not typically meant to map insightful experiences, cognitive and emotional aspects when people interact with and evaluate products or ideas (Cascini et al., 2013). This led to the streamlining of a specific questionnaire to conduct the experiment. Questionnaires allow more people to answer at the same time, thus reducing the time required to acquire the information which needs to be processed. Other relevant strengths include the fact that they are anonymous, the opportunity to "compare responses across various groups and products", and the contribution of digital/online tools for the organisation of data (Wilson, 2013). The importance of the latter point is also underlined by Dinar et al. (2015), who claim the need to shift towards increasingly automated systems when dealing with information relevant to design. Weaknesses in administering questionnaires are likewise acknowledged, but they do not prevent design scholars from using them for exploratory studies about product evaluation. Still with reference to Wilson (2013), the main drawbacks are the skills required to administer and iterations to streamline questionnaires, and retrospective phenomena. As for the former, the authors have gone to considerable effort to define suitable questions and sequences thereof (Section 2.2). The retrospective phenomena are considered poorly relevant here, as respondents are expected to comment on interactions taking place shortly before answering the questions posed.

2. Development of the experimental protocol

The arena of potential users of surprising products is as large as the market and therefore different individuals might be surprised in different ways by the same design. Moreover, different individuals might have different reactions, since one individual might be surprised while another one is not. As the reactions might vary, unveiling the nature of surprise emergence requires the investigation of large datasets to increase the chances of capturing nuances of the phenomenon. The understanding of these different nuances might provide wider opportunities for the development of methods and tools to support design for surprise.

The sources of unexpectedness of a product might reside in what it is, what it is meant for and how it interacts (Desmet, 2012). This leads to the identification of the FBS model as an appropriate candidate to characterise surprise caused by products since it leverages three kinds of variables to describe product distinctive features comprehensively. Such descriptions can be therefore generalised by FBS variables to define design strategies designers can fruitfully reuse. Hence, within the overall purpose of investigating the emergence of surprise in users of products, the authors consistently developed the following experimental protocol with a twofold objective:

- To check the appropriateness and the robustness of an FBS-based cognitive framework (Gero and Kannengiesser, 2004), shifted to the user’s perspective, in describing the processes of surprise emergence through the validity of its constructs;

- To define a repeatable approach to gather large experimental data related to surprise emergence, so as to ensure the reliability of results while limiting the human resources involved.

The following subsections present the original framework developed by the authors that describes design-relevant variables and cognitive processes involved during surprise emergence from the perspective of a user (observer) of a design. Then, as large datasets typically require for a significant amount of data processing, the authors opted to gather data as subjects' answers in a written format in order to minimise the time to transcribe protocols or interviews. The gathering, coding and analysis of the dataset are described with reference to the abovementioned objectives. The whole procedure described in the paper is depicted in Figure 1.

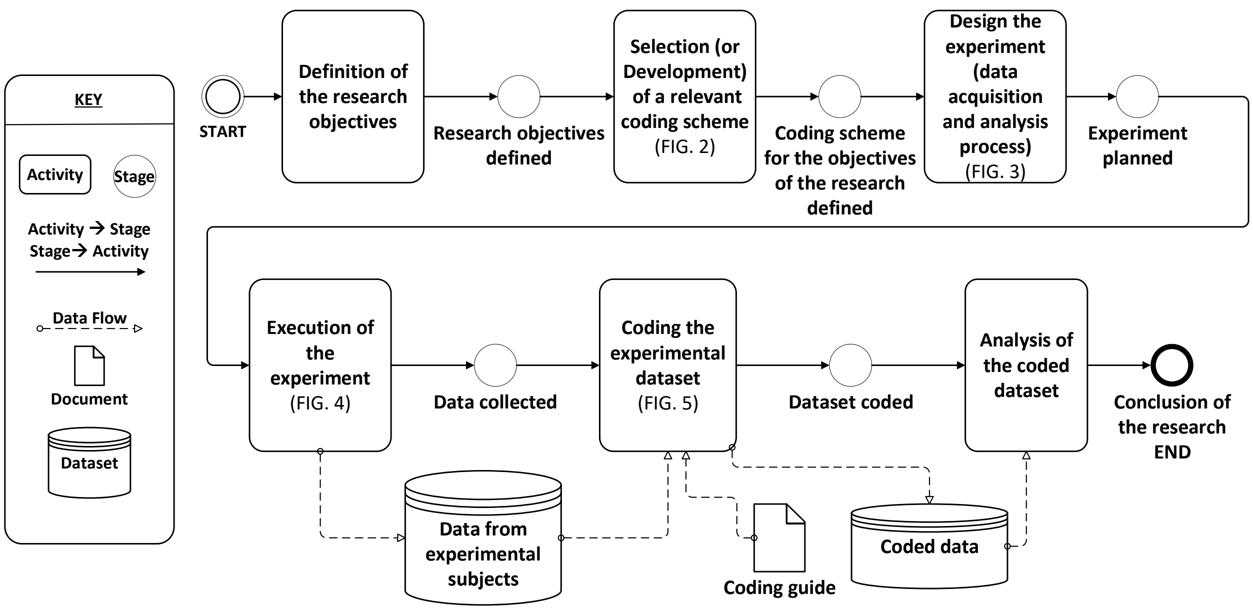


Figure 1. Development of the experimental protocol to analyse surprise emergence in a large number of product evaluators

2.1 Coding scheme: a framework describing the cognitive processes involved in surprise emergence

The authors selected the situated FBS framework (Gero and Kannengiesser, 2004) as the reference to describe the process of surprise emergence by means of design variables characterising product features. Those are also the variables designers typically handle, which facilitates the easier interpretation of surprise triggers. The original FBS ontology (Gero and Rosenman, 1990) and the situated FBS framework, indeed, describe designers' cognition or their thinking processes when they create new operational solutions. The process of surprise emergence is seemingly a situated phenomenon as well, although from a completely different perspective, i.e. from the viewpoint of the product user. Despite the fact that cognitive processes for designers and users can be significantly different as an overall sequence, the ones characterising the interpretation and the evaluation of solutions by designers can be fruitfully reused to describe the users' cognition. Indeed, previous studies from Becattini et al. (2015; 2017a) have already shown the effectiveness of the FBS ontology and the three worlds of situatedness in characterising surprise as the mismatch between the interpretation of the external world and internally preconceived expectations (see Figure 2). Here, the subscripts *e* and *i* stand for external and interpreted, respectively, while the superscript *e* means expected. For clarity, the three worlds have been depicted separately in the illustration, although the expected world should be formally considered as part of the interpreted world. This clarifies that the variables in the expected world exist before the current interpretation that leads to surprise emergence (X_i^e exist in the users' mind before they interpret something new which they are exposed to). The users, indeed, create these expectations according to their previous experience.

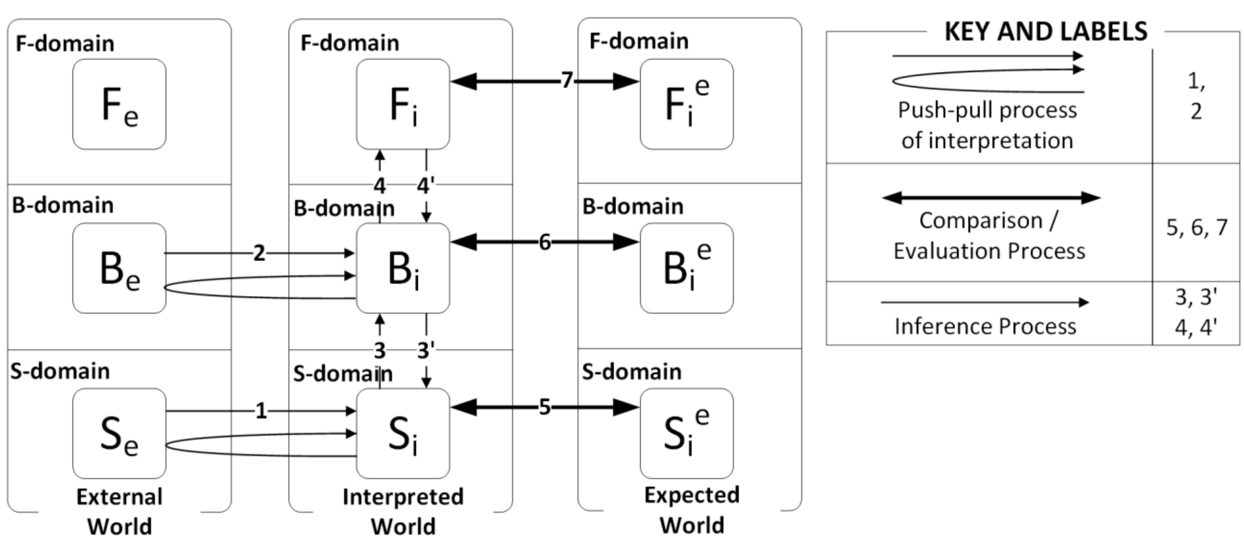


Figure 2. The framework to describe the emergence of surprise as a situated phenomenon and the key to the cognitive processes considered here

The framework describes the emergence of surprise as a mismatch between what is sensed and what is expected and this may occur across different FBS domains or in a single one, where FBS are variables characterising different aspects of the design object:

- F - Function (what is the design for, what is its purpose);
- B - Behaviour (how does the design actually work [Bs] or is it expected to work [Be]); and
- S - Structure (what is the design made of, what is it).

With respect to Gero's and Kannengiesser's (2004) framework, the three worlds assume the following meaning in light of the shift of perspective from the designer to the user.

- "The *external world* is composed of representations outside the <user>";
- "The *interpreted world* is built up inside the <user> in terms of sensory experiences, percepts and concepts";
- "The *expected world* "holds the figured-out projections of what was previously experienced".

Such users' expectations belong to their constructive memory and allow them to arrange the interpreted meaning of what is sensed in the external world. Constructive memory, indeed, is a process by which one builds up meaning about what is sensed and perceived. One digs into one's own knowledge in a push-pull process — the detailed dynamics are described in Kelly and Gero (2014). The push process is initially driven by the data sensed in the external world and the pull process is driven by previous or updated expectations that determine the way someone interprets what is sensed. Indeed, this process occurs between the external and the interpreted worlds, as well as in the interpreted world itself as a result of knowledge retrieval from previously conceived expectations.

Beyond the concepts from Gero and Kannengiesser (2004), the framework also postulates additional cognitive processes involved in surprise emergence. They occur in sequence across the three different worlds. First, between the external and the interpreted world; then, between the interpreted and the expected world. Inferences in the world of interpretation can be assumed to be natural ways to link design variables, e.g. by making hypotheses, assumptions or more simple associations. The following paragraphs describe the process-linking variables between worlds and within the interpreted world.

First, push-pull processes (1 and 2 in Figure 2) prompt users to scan their memory and build meaning out of what is sensed in the external world in order to interpret it. These processes, as for Gero's situated framework, occur across different worlds, but within the same domain as the FBS ontology. In addition, the authors postulate here that push-pull processes of interpretation can occur in the FBS domains where senses operate; therefore, functions [F], as abstract representations of a product's purposes, cannot be sensed. Their interpretation requires the second and next cognitive process.

Second, within the interpreted world (as well as in the expected world), a different kind of cognitive process occurs for the conceptualisation of what is sensed and perceived, in order to interpret it thoroughly. This kind of process occurs when the user creates a link between two distinct design variables belonging to the same product. In other words, what is sensed with reference to a certain FBS ontological variable can trigger reasoning by deductive inference (Roozenburg and Eekels, 1995) in order to define different kinds of variables for the same sensed entity. For instance, seeing or touching a hole in a water basin [S] might make the subject suppose the water will pour out [B]. Therefore, these inference processes (3 and 4 in Figure 2) occur across different FBS domains and within the same world. Moreover, there is no preferential direction for deductive reasoning between variables.

Third, the interpretation is checked with previously conceived expectations for the same FBS ontological variable between the interpreted and the expected word. The proposed framework suggests adopting a comparison (5, 6 and 7 in Figure 2), as the cognitive process responsible for surprise emergence, between pre-existing expectations and what is sensed and conceptualised. This comparison is in line with Gero's framework, despite originally simply describing the designers' comparison between the expected behaviour [Be] and that of the actual structure

[Bs]. However, given the perspective change from a designer to a user (here an observing agent), the comparison can also include structural [S] or functional variables [F], e.g. miniaturised devices might surprise those who have just experienced standard-sized ones.

The above elements constitute the core concepts of the proposed framework; they also serve as the fundamentals of a coding scheme to interpret the answers of subjects involved in the experiment. Because of the potential bias resulting from questions that trigger cognition towards framework-dependent paths, the authors carefully formulated questions and their sequence, as described in the next subsection.

2.2. Plan for the experiment: development of the questionnaire

The final form of the questionnaire is the result of a three-step iterative process, summarised in Figure 3, during which the effectiveness of the questions and their robustness beyond language barriers have been checked and refined. In general, the initial questions attempt to extract the main reasons behind surprise emergence. Moreover, in all versions, some questions specifically target a single FBS variable each, in order to give a complete picture of surprise triggers. However, the formulation of these questions uses non-specialist jargon and aims to verify whether subjects would have identified, in their own words, that certain variables in question were responsible for triggering surprise.

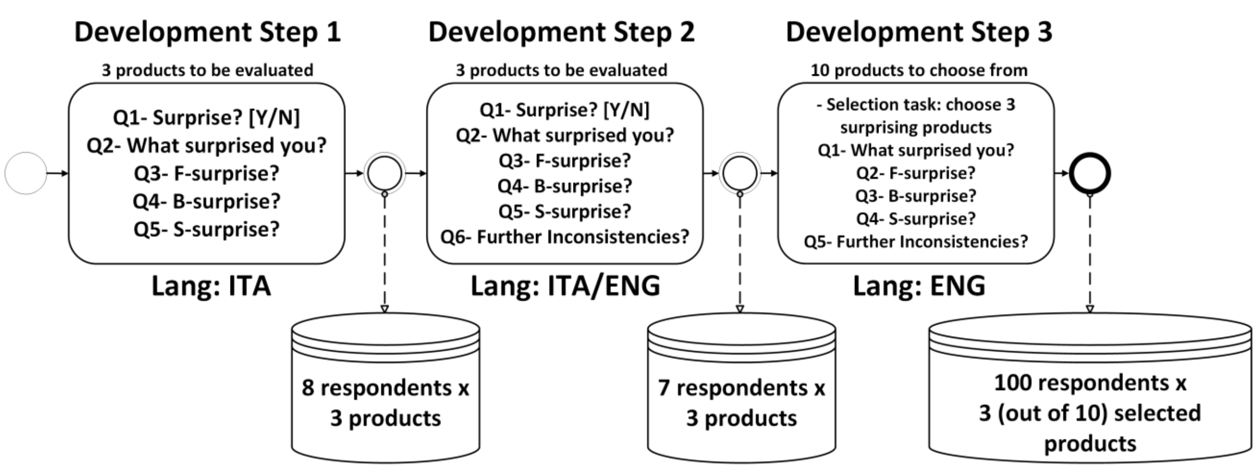
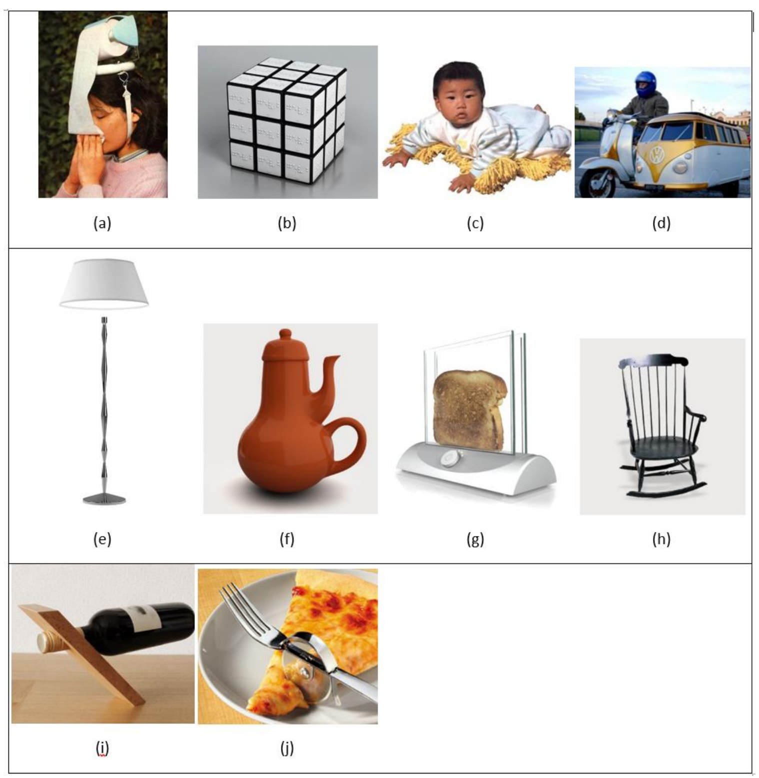


Figure 3. Description of the process used to refine the questionnaire included in the experimental protocol (continuous and dashed lines as for the key in Figure 2)

In the final version, each question indicates the expected number of words for the corresponding answer, because the excessive concision of a few answers in previous rounds prevented the mapping of inherent cognition. The final version allows respondents to choose the three most surprising products out of a set of ten items proposed through on-screen pictures. This measure was deemed necessary to allow the subjects to focus on the products which engendered the most surprise. The subjects also received some initial instructions on how to answer the questionnaire. The following text juxtaposed with Figure 4 authentically reflects the administered questionnaire.

> *Here you find pictures of 10 products, labelled with letters from (a) to (j), that are supposed to be surprising. Without considering your judgement about the aesthetics of the product, which is not interesting to the purpose of the present survey,*[1] *select three of them which you judge the most surprising. For any of them answer the following questions, by specifying the letter associated with the reference picture.*



[1] This recommendation was given to avoid answers justifying selection of products based on personal preferences.

Figure 4. The set of ten objects from which respondents had to select the three they found most surprising

1. *What are the product characteristics/features that you did not expect and were essential to surprise you? (10-50 words)*
2. *Is it clear what the product is for? If yes, explicate it; if not, explain what is not clear and why it is not clear. (20-100 words)*
3. *Can you clearly understand how the product works? If yes, describe your understanding (20-100 words); if it is not clear how it works, please explain what is not clear and why it is not clear. (20-100 words)*
4. *Are there characteristics or features that you are not used to perceiving and that you did not expect in this context? If yes, please explain which ones and why. (20-100 words)*
5. *More generally, did you find any inconsistency between the product features, the way it works and its purpose? Please explain which ones and make explicit if these inconsistencies triggered surprise to some extent. (20-100 words)*

2.3 Execution of the experiment, gathering and coding of the dataset

The questionnaire was developed in a digital format to facilitate sharing and to help in the automation of answer management. It was administered to 100 anonymous and randomly selected subjects (potential users of the solutions in Figure 4). They were unaware of the scope of the study and their knowledge of design topics and the FBS was considered inessential. They were hired through http://www.microworkers.com, a web platform that gathers a community of workers who receive a reasonable salary for small tasks they choose to carry out — each questionnaire which was properly filled out was rewarded with 3 USD.

The questionnaire could be completed in 45 minutes or less, but subjects were allowed to take longer. The web service collects the input data as text. Results are provided as comma-separated values and each row corresponds to the answers of one single subject.

Answers then require coding according to the original framework proposed above. To this purpose, coders (among the authors in this study) had to process the results by interpreting them and identifying clear references to the FBS ontology, and the worlds they belong to. The coding should reveal which ontological variables are involved in surprise emergence and how they are linked to subjects' inherent reasoning. More precisely, to check the effectiveness of the proposed framework, the interpretation and coding of the provided answers should be capable of identifying:

1) The FBS ontological variables that apparently determine surprise emergence
2) The cognitive processes occurring between two FBS variables
    - of the same kind between different worlds (processes 1, 2, 5, 6, 7 in Figure 2); and
    - of different kinds within the same world (processes 3, 3', 4, 4' in Figure 2).

Shifts between FBS variables within the same world (S←→B or B←→F) can, however, be simply inferred, as questions targeting these shifts would presumably affect cognition during product evaluation. Hence, these processes will be generically referred to 3 and 4, respectively.

The interpretation of results introduces a further (communication) bias beyond the potential misalignment between the subjects' cognition and what they report. To minimise such a bias, a sufficiently large number of independent coders analysed the answers individually.

Since the approach and the coding scheme had not been used before, they needed to be consolidated. As illustrated in Figure 5, the coders first analysed a pilot dataset of randomly selected answers (Activity 1.1 – sub-dataset A: 20% → 60 answers). This analysis (Activity 1.2) supported the development of a coding guide (supplementary material), which addressed emerging comments and difficulties — its contents benefitted from insightful design knowledge offered by Vermaas (2009) and Gero and Kannengiesser (2004). The coding guide led to an approximate doubling of IRR score; details follow. All the answers were considered in the final stage, as made explicit through Activities 2.1 (coding of sub-dataset B: 80% → 240 answers) and 2.2 (recoding of sub-dataset A). The coders also noted the time devoted to each stage to monitor human efforts.

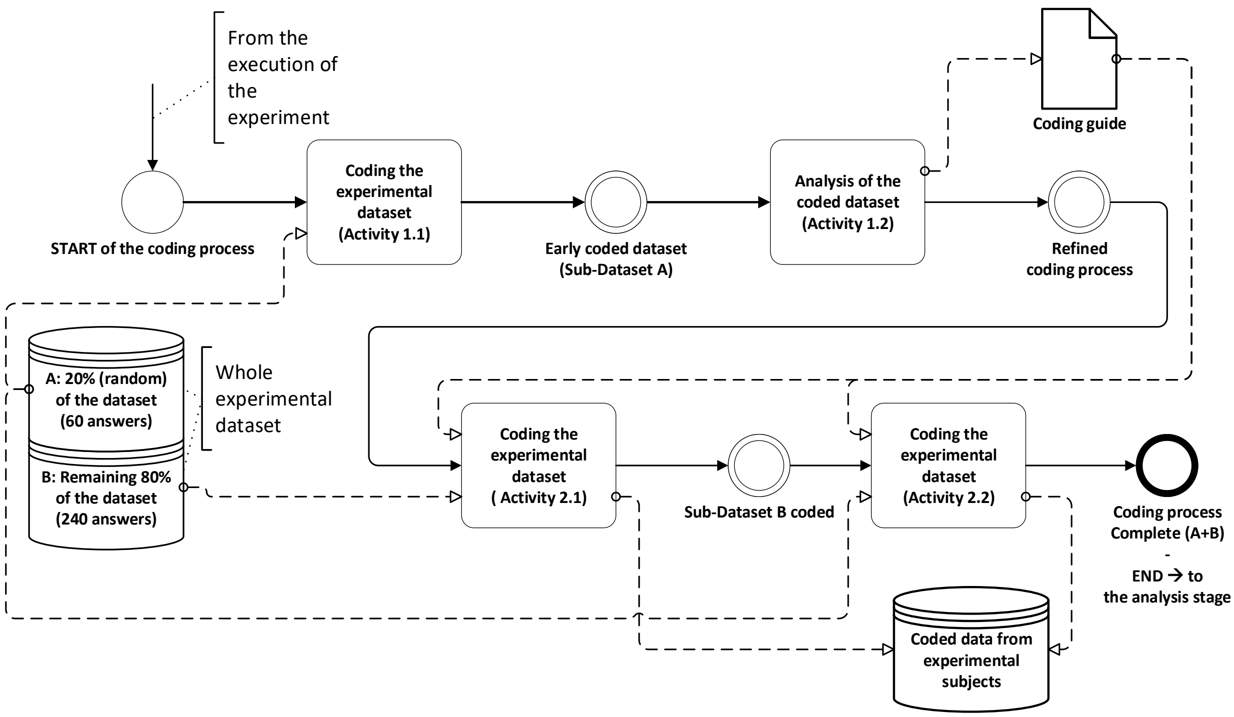


Figure 5. The refinement of the coding procedure that highlights the introduction of a coding scheme to establish formalised criteria for the interpretation of respondents' answers (continuous and dashed lines as for the legend in Figure 1)

2.4 Analysis of the coded dataset

Agreement among coders represented a direct measurement of the reliability of the analysis and the possibility of extracting meaningful data. It also enabled an indirect check of the robustness of the coding scheme to provide a comprehensive opportunity to map for the concepts/constructs of interest, the appropriateness of descriptions, and the absence of major omissions (Snider, 2014).

The level of agreement between coders was measured by Inter-Rater Reliability (IRR) and expressed through statistical coefficients. Consistently with the coded constructs, which are recalled in the numbered list in the previous subsection, the Fleiss' Kappa coefficient was used in the present study.

The subsequent interpretation of IRR data was in line with the rule of thumb introduced in Landis and Koch (1977), for which a substantial agreement (Kappa > 0.6) is often set as a target to validate models by means of coding convergence.

3. Subjects' answers and outcomes of the coding process

The robustness, viability and effectiveness of the proposed approach are here presented together with the results of the investigation. Section 3.1 presents some illustrative answers characterising surprise emergence and different coding interpretations. Section 3.2 presents the analysis of the IRR for the constructs indicated in 2.4. In Section 3.3, outcomes in terms of human effort are reported. Section 3.4 identifies the information coders could extract from the answers of surprised subjects and explains how the coding scheme is capable of distinguishing different modalities of surprise emergence.

3.1 FBS variables and cognitive processes

Table 1 presents five answers selected from the set of 300 available replies (3 products for the 100 subjects who participated in the investigation). These answers were selected to demonstrate the commonalities and differences the approach spotlights and the ways coders classified them.

Table 1: Sample of answers according to the proposed questionnaire and corresponding coders' interpretations

Product ID	Answers	Coder 1		Coder 2		Coder 3	
		FBS var.	Processes	FBS var.	Processes	FBS var.	Processes
A	What surprised you?: I have never seen someone with toilet paper on their head F-SURPRISE?: I guess for blowing your nose but otherwise I do not know B-SURPRISE?: Pull the toilet paper down off the roll, wipe your nose, sits on your head so I think I can understand how it works S-SURPRISE?: Yes toilet paper on somebody's head is not something I have seen or thought of before why not just a tissue FURTHER INCONSISTENCIES?: Toilet paper being used not tissues and on top of paper is usually used for another purpose	S, F	1, 3, 4, 5, 7	S, F	1, 3, 4, 5, 7	S	1, 3, 4, 5, 7
C	WHAT SURPRISED YOU?: I wasn't expecting to see a baby gro with a floor polisher attached to it. F-SURPRISE?: Yes, totally clear what the product is for - it's so the baby can polish the floor as it crawls around. B-SURPRISE?: Yes, I can understand how the product works - baby crawls across floor, and as it moves the polishing 'fronds' buff the floor at the same time. (Ingenious!) S-SURPRISE?: I've never seen anything like this before. I've seen baby gros before, and floor polishing cloths, but never the two combined in one baby gro. FURTHER INCONSISTENCIES?: I didn't notice any inconsistencies between the product and the way it works and it's purpose. I'm not sure how good it would be at actually polishing though. Babies don't crawl for very long, and they'll soon be cruising and pulling themselves up, so a very limited use product. Looks like it might also be a dirt trap.	S	1, 3, 4, 5	S, F	1, 5, 3, 4, 7	S	1, 4, 3, 5
E	WHAT SURPRISED YOU?: I found the fact that the lamp cover is hovering above the post to be surprising since I thought that the bulb had to be attached to the post for the lamp to function F-SURPRISE?: It is clear by the design that this is a lamp, used to light a specific area. It has the same design and features as many other lamps, except for the fact that the cover hovers above the post. B-SURPRISE?: It is not clear how this works because I don't know how the lamp would light the room if the cover/bulb are not attached to the lamp post. It may be controlled wirelessly, but I am not sure by just looking at the picture. S-SURPRISE?: Yes, I was not expecting the lamp cover to be hovering above the post; it is not clear how the lamp works because of this. Normally the lamp cover and bulb are attached to the post. FURTHER INCONSISTENCIES?: Yes, it seems inconsistent the lamp cover and bulb would be hovering above the post. One would think that the cover and bulb would need to be attached to the post in order to light/turn on, so it is surprising to see that they are not.	S, B	1, 3, 4, 5, 6	B, S	1, 3, 4, 5, 6	B, S	1, 3, 4, 5, 6
I	WHAT SURPRISED YOU?: The wine bottle seems to be defying gravity. It takes a second to realize what's happening. F-SURPRISE?: Yes. It's a very cool looking wine bottle holder that plays a trick on the eyes. It would be a great conversation piece. B-SURPRISE?: Yes. There really isn't a whole lot to explain. It's a wine bottle holder that appears to be delicately balanced to give the illusion that it's defying gravity. S-SURPRISE?: Yes. It plays a trick on the eyes when you first see it because it looks like the bottle should topple over but it doesn't. FURTHER INCONSISTENCIES?: Not really. It seems perfectly straightforward to me. It's just a fancy wine holder that gives the illusion defying gravity.	B	1, 3, 4, 6	B	1, 3, 4, 6	B	1, 3, 4, 6
J	WHAT SURPRISED YOU?: The pizza cutter and fork are one unit. Usually they are separate utensils. F-SURPRISE?: Cut the pizza with the circular cutter, then use the fork to feed the cut off piece to one's mouth. It's designed to save time and motion. B-SURPRISE?: Cut the pizza by rolling the circular cutter over the pizza. Then use the fork to eat it. It can be done with one hand. S-SURPRISE?: Usually kitchen utensils have only one function. This has two. And it can be done with only one hand. Saves motion and time. I can eat faster. FURTHER INCONSISTENCIES?: Usually it's a 2-step process to cut up a pizza and then eat it. And the cutting tool and eating tool are usually separate. This product combines the two. Highly functional and time-saving. The older generation might frown on it though.	S, F	1, 3, 4, 5, 7	S, F	1, 3, 4, 5, 7	S, F	1, 3, 4, 5, 7

3.2 Inter-rater reliability

IRR has been analysed according to the indications included in Section 2.4. It refers to the final version of the coding (stage 2, Figure 4). Consistent with common practice in design research (e.g. Srinivasan et al., 2018), three authors coded the dataset. FBS combinations were treated as nominal variables, giving rise to eight different outputs the coders could choose from (including the absence of ontological variables). Therefore, 300 cases (3 products for each respondent) are available for the evaluation of the IRR. The extremely large number of combinations of processes (1, 2, 3, 4, 5, 6, 7 in Figure 2) led the authors to consider whether to assess the absence/presence of each process through dummy variables. The calculation of the Kappa coefficient reflects 2100 classifications of the dummy variable, by considering the seven coded processes in the 300 cases for this study.

The Fleiss' Kappa results indicate a substantial agreement for FBS variables triggering surprise emergence (IRR=0.64) and for inherent situated FBS cognitive processes (IRR=0.75). This indicates that the constructs of the proposed framework are sufficiently robust and they lead to converging interpretation of modalities of surprise emergence.

In addition, only a few cases (details in 3.4) could not be ascribed to FBS variables, again indicating that the whole procedure can be considered to have been robust.

3.3 Human effort

The following bullet list reports the human effort involved in coding each set of answers. The information provided about average time does not differentiate between learning time and the regime condition, which is supposedly considerably smaller than the indicated effort.

- First step (Activity 1.1 in Figure 5), 20% of randomly selected answers: approximately 3'30 per coder per surprising product.
- Second step (Activities 2.1 and 2.2 in Figure 5), 100% of answers: approximately 1'30 per coder per surprising product.

The resources used to finalise the coding guide – which was deemed necessary due to several doubts arising among coders – totalled approximately four hours per author. In the event of the experiment being repeated by other scholars within the field of surprising designs, these resources would not be required.

3.4 Insights into modalities of surprise emergence

The robustness of the investigation protocol allows some preliminary comments to be made on the phenomena linked with surprise emergence.

The example of Table 2, which details two answers for the Rubik's cube for blind people, clarifies that the same product can trigger different reactions of surprise as they (mostly) relate to two different variables in the FBS ontology.

Table 2. Illustrative example of a product that triggered surprise in different modalities

	Surprise due to Structural Variables	Surprise due to Functional Variables
What surprised you?	When you see a product like this, you expect a multi-colored Rubix cube. This, however, is all white.	I think the idea of the blind being able to do a Rubik's cube is incredible. I think it is a nice spin on an old classic
F-surprise?	No, it is not clear what this is for. A typical cube like this is a game in which you match colors. Without this having colors, it makes no sense what it is for.	Yes, I believe it is very clear that it is a Rubik's cube for the blind. I think that anyone with knowledge of how the blind read, would know this immediately.
B-surprise?	I do not personally understand how this product works or what the intentions of the manufacturers would be. Maybe there is more to it that cannot be seen from the picture, or it is something used in a foreign country, but as is I do not see how this could possibly work for anything.	Yes. It is a basic Rubik's cube just with font on the squares instead of color. It would be difficult maybe at first for a blind person to use it at first, but I feel it would be easy to get use to. It is just hard to solve the puzzle.
S-surprise?	Yes, I am not particularly used to perceiving an all white Rubix cube. When you see one of these, you expect to see many colors and understand that it is a game.	I am not use to seeing braille at all. I don't know too many people who are blind. I feel that the complete white color of a Rubik's cube throws me off as well. It should still probably have color to it.
Further Inconsistencies?	The major inconsistency in this product is the lack of colored squares. This defeats its purpose, as you can not play the game as intended.	I think the product should be colored so it could be a combined effort between the blind and somebody who would want to help. Unless you read braille and you could see, you wouldn't be able to help.

The answer to “what surprised you?” in the left-hand column claims that surprise is due to the colour of the Rubik’s cube faces (structural variable). In turn, the other respondent (right column) is surprised by the idea that a product allows blind people to play with a Rubik’s cube, which concerns the purpose of the product (functional variable). The other answers also show that the emergence of surprise strongly depends on the user and on the proper perception of the product: the subject surprised by the colour of the Rubik’s cube faces did not notice any Braille characters, while the other subject did. These answers also demonstrate that the phenomenon of surprise can also occur due to different cognitive process chains. Both the subjects observed the product’s structural variable – Push-pull process 1 in Figure 2 – as the modality of administration of the questionnaire through static images reduced the chance of perceiving the product behaviour directly. Both the subjects also strove to find a meaning for the purpose of the product and how it should behave (inference processes 3 and 4 in Figure 2).

The answers collected in the left-hand column demonstrate that failing to perceive the braille characters hampers a correct understanding of both the behaviour and the purpose of the product. This also confirms that the subject found the colour of the faces of the Rubik's cube surprising, as he/she compared it with the coloured faces of the traditional product (comparison process 5 in Figure 2). The subject's answers in the right-hand column show that a full understanding of the product structure does not result in surprise at a structural level, but due to a different design variable. The appropriate interpretation of the product's purpose allows for a check against the related expected functional variable (comparison process 7 in Figure 2), where expectations are clearly violated. Similarly, but for a different product, the authors collected answers about the surprising behaviour of a side-rocking chair: the simple tilted placement of the rockers was not considered surprising, but the inferred behaviour was (comparison process 6 in Figure 2).

Despite the fact that emotions and the characterisation of positive and negative surprise go beyond the purpose of this paper, the two sets of answers in Table 2 also provide evidence that FBS variables are appropriate for describing both kinds of reactions.

Figure 6 shows the distribution of the 300 answers as they emerged from the activity of the coders. FBS variables are here designated upon full agreement of at least two coders. The Figure distinguishes when the absence of FBS variables is caused by:

- the impossibility of reaching agreement between coders ("Not univocal"), i.e. the phenomenon of surprise emergence is described by FBS variables, but not univocally determined by at least two of them; and
- a lack of clarity or absence of information about the reasons for surprise emergence ("Off topic"), e.g. the respondent discusses the product, but not how surprise is triggered.

As is apparent from Figure 6, the experimental protocol allowed the identification of variables involved in the emergence of surprise in more than 90% of the coded answers, as just 9.7% of them resulted in ambiguous interpretations of FBS variables. The figure also shows that only 5% of answers did not address the question properly.

The graph shows a strong prevalence for S-driven surprise emergence. This could be ascribed to the modality of questionnaire administration, as static visual content (photos, pictures) mostly stimulate spatial perception.

Couples of variables (FS, FB, SB) are involved in a small fraction of answers, approximately 1/8 of the whole set, while no couple of coders individuated answers spanning the whole set of FBS variables in a single instance. Overall, S-driven surprise emergence is strongly prevailing here too, as only two out of the 300 answers combine F and B variables.

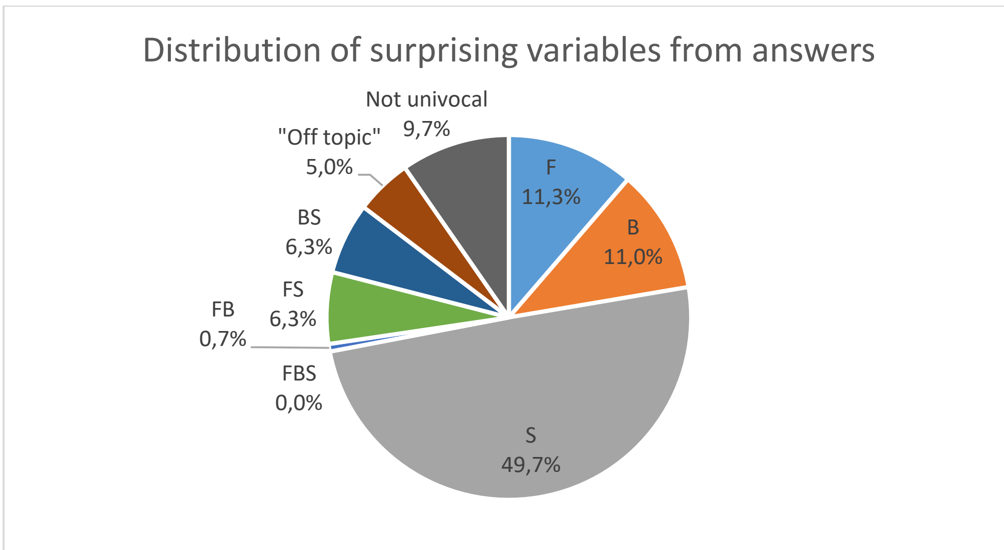


Figure 6: distribution of coding results in terms of FBS variables responsible for the emergence of surprise

Figure 7 depicts the number of occurrences of cognitive processes in the 300 answers as interpreted by the coders. The graph shows that the modalities of product interpretation are (almost) exclusively structure-driven (process 1), which is consistent with the above considerations on the presentation of product forms. It is, indeed, worth noting that any direct interpretation of product behaviour (process 2) from static images was very rare, as Be→Bi processes were noted in only two out of 300 answers. Deduction processes that link variables in the interpreted world (Si→Bi – 3 and 3'; Bi→Fi – 4 and 4') occur extremely frequently, in a way which is comparable with the push-pull processes of interpretation from the external world. The nature of the administered questions might bias the observers' natural flow of thinking, leading them towards partially induced reflections and deductive reasoning among FBS variables. More open-ended questions would probably reduce the effect of such a bias, but significantly slow down the coding of answers. The cognitive processes directly involved in processes of surprise emergence are supposedly those for which a subject compares the same design variable between two worlds: the interpreted and the expected ones. The mismatch between the interpreted variable and its expected corresponding one leads to surprise emergence according to the framework proposed in Figure 2. The right part of Figure 7 shows that the process of comparing two variables occurs very frequently (processes 5, 6 and 7) and, in some cases, it involves more than one kind of variable; this is also suggested by the modalities of surprise emergence shown in Figure 6.

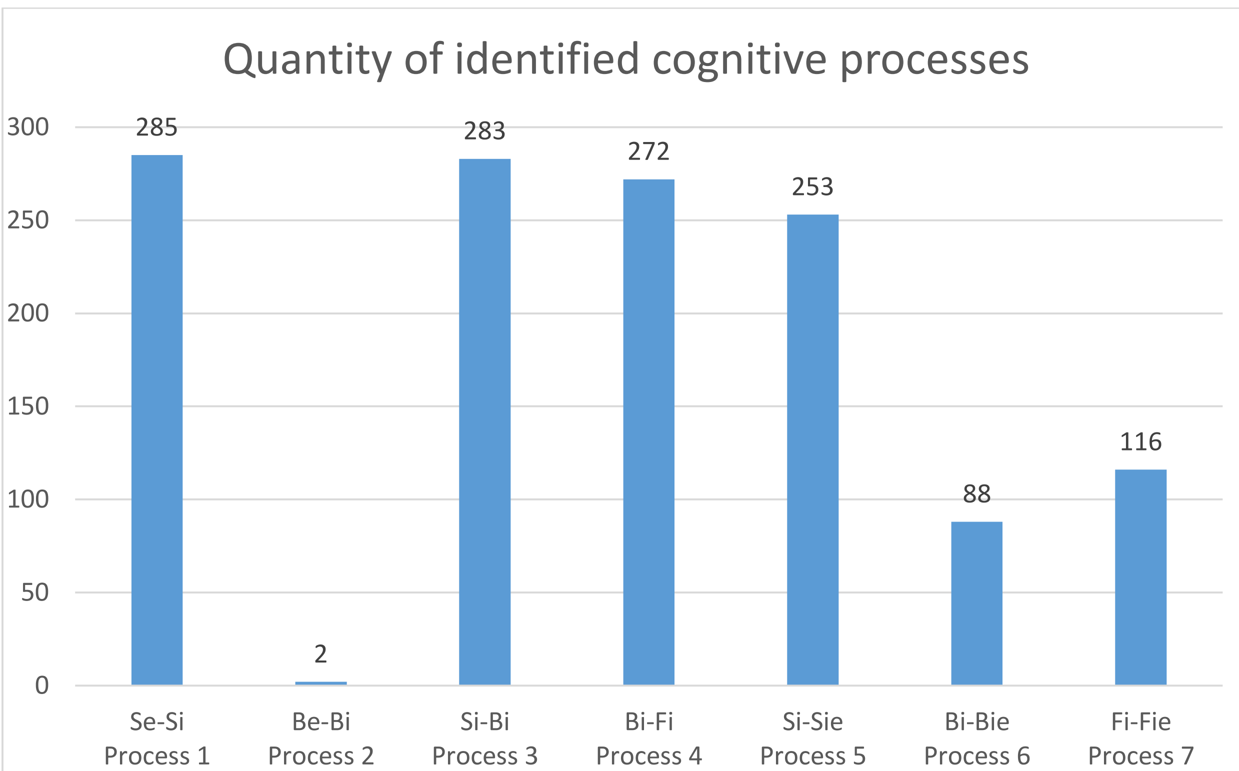


Figure 7: Distribution of coding results in terms of mapped processes of surprise emergence. The graph recalls the situated variables involved in the processes; acronyms and process numbers are as for Figure 2

4. Discussion

4.1 Main achievements and limitations

The developed protocol has enabled a large number of answers provided by non-designers (potential users) relating to their reactions to surprising products to be gathered and analysed. The experiment has verified the suitability of pre-existing design ontologies (FBS variables) to explain surprise phenomena. In this respect, the twofold objective illustrated in Section 2 has been achieved, though some limitations have to be acknowledged (see below).

A tailored questionnaire, avoiding design-specific jargon, highlighted the main design variables triggering the emergence of surprise. This approach was successfully in light of the limited number of answers that did not include the expected information (Figure 6). This paper supports the possibility of using targeted questionnaires for design research despite their inherent limitations, as already underlined in Section 1.4. The research confirms that questionnaires are advisable when large populations need to be involved (Crilly and Cardoso, 2017).

The effort required in the research process, and markedly in post-processing, is often reported as relevant to choosing the methods for acquiring design knowledge. Strengths and weaknesses of the authors' approach are here presented. With regard to data acquisition, the approach required minimal time for administering and processing, especially given that the questionnaire was submitted online. This is in addition to the fact that questionnaires allow for simultaneous multiple pen-and-paper completions. In this regard, it is worth remarking that the answers from 100 people were gathered in just a few days and the cost in terms of monetary rewards for participants was low.

The involvement of anonymous subjects with likely no knowledge about design research and FBS was a requirement of the study. This may have been detrimental in one sense, as it was not possible to control for these subjects' commitment and reliability. The validity of the study was safeguarded by the large number of subjects, however, which may have compensated for possible inaccuracies in some of the answers, and the satisfaction of IRR values, which demonstrate a clear convergence towards the coded aspects. Therefore, the present research supports the possibility of using crowdsourcing platforms to obtain large numbers of answers.

Moreover, although the development of a bespoke questionnaire was undertaken carefully, it is not possible to extract every potentially relevant piece of information with certainty, and, additionally, extracted information might depend on the formulation of questions. Missing statements may, potentially, have affected the dataset too. In addition, the option to select three products to focus on clearly enabled more information to be gathered (300 artefacts instead of 100), but may have introduced a bias. Subjects already acquainted with the questionnaire may have changed their answering strategy for products 2 and 3, i.e. once they knew the questions. Nevertheless, a check on the first answers confirms that this effect is almost negligible: the FBS variables triggering surprise in the first answers are distributed consistently with the whole set of 300 answers (Pearson correlation: 0.98). In order to confirm this conclusion, a case study involving 300 participants who select a single surprising product could be undertaken to rule out the presence of this kind of bias.

As for coding information, the proposed approach minimised the need to transcribe and segment the answers in comparison to the work required to process answers from the unstructured text of interviews. No advantages or disadvantages are to be claimed with respect to the analysis of the information, as this depends on the criteria used. The time spent by coders, especially after the development of the coding guide, appears to have been acceptable. In this sense, the use of an acknowledged framework for the coding scheme is supposedly supportive of the learning process for coders. Although three authors worked as coders, it is expected that the approach would be suitable for replication by external experts, who would benefit from the existing coding guide. With respect to this, however, new coders may require an introduction to the FBS-based coding scheme in consideration of the shift of perspective from designers to users.

4.2 Applicability of the protocol and significance of the preliminary outcomes

The ability to explore and characterise surprise emergence with the available data is still questionable, given the limited number of subjects and products investigated. Moreover, the employment of some intentionally bizarre products allowed the individuation of at least three pictures which were surprising to all the subjects, but these artefacts could engender different surprise phenomena in the context of surprising designs meant for commercial purposes. Moreover, the initial sample of selected products may have been biased — all the products were deemed somehow surprising by the web sources they were found in, but these interpretations were not backed up by any scientific claims. Nevertheless, the literature shows that a preconceived and ready-to-use set of surprising products to elicit users' reactions is lacking and the experimental outcomes confirm that all ten products were surprising, since each of them was chosen more than once. In this context, users were asked to comment on surprise with respect to the products depicted in the images, (which resulted in surprise due to structural features, as documented above). It is hypothesised that different reactions may take place when different forms of interaction are leveraged, for instance by means of text, sound, video, physical prototypes, virtual reality and real-life experiences, as for the findings presented by Berni et al. (2020). The opportunity to perform tests with large numbers of subjects, with altered illustration media, depends on the resources and the constraints posed by online systems, if employed. Nevertheless, those changes are not expected to challenge the suitability of the FBS framework for describing surprise emergence. Therefore, in many respects, the conclusions drawn in Section 3.4 might be considered partial and limited by the conditions of the experiment, since the modality of questionnaire administration did not allow for a comprehensive exploration of alternatives. In fact, the way surprise emerges can also vary significantly according to the nature of images (not just photographs, but also sketches, rendering, 3D models…). On the other hand, the described limitations did not hinder the validity of the criteria and the protocol for studying surprise, which consequently represents a candidate benchmark for future studies in the field. It should be noted that the chosen FBS ontology enables the description of product features responsible for surprise emergence, but the existence of other suitable taxonomies and coding schemes cannot be excluded. However, the framework employed for this study offers a key advantage in that it includes concepts many designers are already familiar with. Indeed, information exchange from the user to the designer domain can take place smoothly once FBS-based "design for surprise" guidelines have been developed (by, for instance, replicating the surprising effects described in the answers presented in Table 2).

In addition, the authors believe that the testing protocol has a repeatable and versatile structure for the execution of tests where users (non-designers) have to evaluate different products, ideas and concepts, even outside the surprise domain.

As has already been suggested, different features and variables can be monitored by means of results extracted despite the fact that the questionnaire had a particular focus. For instance, abstract distinctions for unexpected structures are worth investigating given the large number of reflections on structural aspects. Moreover, the positive or negative valence of surprise might surface from the provided answers. Nevertheless, to allow this kind of finding to emerge, a tailored selection of products is crucial to draw meaningful conclusions.

5. Conclusions

The paper stems from the idea that a better understanding of surprise might represent a turning point in creative product development. This methodological paper argues that, at present, it is necessary to define a cognitive framework underpinning the processes of surprise emergence and benefitting from constructs familiar to designers. While consolidated results in the study of surprise emergence can be ascribed in line with appraisal theory, the research focus has seldom been on products; additionally, the requirements to provoke unexpectedness, as identified, cannot easily be translated into indications useful to designers. The main achievement of this study is in its elaboration of the potential use of the acknowledged FBS framework to explain surprise reactions to products.

According to the current knowledge on surprise in design, the authors identified the need for a shift of focus from the designer’s to the user’s cognition; this involves a paradigm shift in the analysis of design-related characteristics. Scholars have been studying designers’ cognition in recent decades with the twofold objective of developing more effective design methods and ensuring the inherent logic of design thinking and the related rationale are replicable. Said paradigm shift allows a similar exploration of cognition through the same variables. On the one hand, the information derived from external users offers a fresh and relevant perspective on what works better and what creates disaffection from a solution. On the other hand, this also enables the exploration of observation logics inside yet-to-be-developed automated design agents, which evaluate products according to human-like reasoning by resorting to similar or extended coding schemes for cognitive processes.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Ahmed, S. (2007). Empirical research in engineering practice. *Journal of design research*, *6*(3), 359-380.

Becattini N., Borgianni Y., Cascini G., Rotini F. (2015), Surprise as a situated phenomenon. In *DS 80-11 Proceedings of the 20th International Conference on Engineering Design (ICED 15)*, Vol 11: Human Behaviour in Design, Design Education; Milan, Italy, 27-30.07.15, 71-80..

Becattini, N., Borgianni, Y., Cascini, G., & Rotini, F. (2017a). Surprise and design creativity: investigating the drivers of unexpectedness. *International Journal of Design Creativity and Innovation*, *5*(1-2), 29-47.

Becattini, N., Cascini, G., & Montagna, F. (2017b). Exploring the Cognitive Dynamics of Product Appreciation. In Design Computing and Cognition'16 (pp. 555-573). Springer, Cham.

Berni, A.; Maccioni, L.; Borgianni, Y. Observing Pictures and Videos of Creative Products: An Eye Tracking Study. Appl. Sci. 2020, 10, 1480.

Borgianni, Y., & Hatcher, G. (2017). Similarities and differences between humorous and surprising products. In DS 87-8 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 8: Human Behaviour in Design, Vancouver, Canada, 21-25.08. 2017, 31-40.

Bucciarelli, L. L. (1988). An ethnographic perspective on engineering design. *Design studies*, *9*(3), 159-168.

Cascini, G., Fantoni, G., & Montagna, F. (2013). Situating needs and requirements in the FBS framework. Design Studies, 34(5), 636-662.

Chowdhury, A., Reddy, S. M., Chakrabarti, D., & Karmakar, S. (2015). Cognitive theories of product emotion and their applications in emotional product design. In *ICoRD'15–Research into Design Across Boundaries Volume 1* (pp. 329-340). Springer, New Delhi.

Crilly, N. (2015). Fixation and creativity in concept development: The attitudes and practices of expert designers. *Design Studies*, *38*, 54-91.

Crilly, N., & Cardoso, C. (2017). Where next for research on fixation, inspiration and creativity in design?. *Design Studies*, *50*, 1-38.

Dankfort, Z., Roos, L., & Gonçalves, M. (2018). Inspiring co-evolution moves and creativity in design teams. In *DS 89: Proceedings of The Fifth International Conference on Design Creativity (ICDC 2018), University of Bath, Bath, UK* (pp. 395-402).

Demir, E., Desmet, P. M., & Hekkert, P. (2009). Appraisal patterns of emotions in human-product interaction. *International Journal of Design*, *3*(2), 41-51.

Desmet, P. M. (2003). A multilayered model of product emotions. *The Design Journal*, *6*(2), 4-13.

Desmet, P. M. (2012). Faces of product pleasure: 25 positive emotions in human-product interactions. *International Journal of Design*, *6*(2), 1-29.

Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Hernandez, N. V. (2015). Empirical studies of designer thinking: past, present, and future. *Journal of Mechanical Design*, *137*(2), 021101.

Ellsworth, P. C., & Smith, C. A. (1988). Shades of joy: Patterns of appraisal differentiating pleasant emotions. *Cognition & Emotion*, *2*(4), 301-331.

Gero, J. S., & Kannengiesser, U. (2004). The situated function–behaviour–structure framework. *Design studies*, *25*(4), 373-391.

Gero, J. S., & Mc Neill, T. (1998). An approach to the analysis of design protocols. *Design studies*, *19*(1), 21-61.

Gero, J. S., & Rosenman, M. A. (1990). A conceptual framework for knowledge-based design research at Sydney University's Design Computing Unit. *Artificial Intelligence in Engineering*, *5*(2), 65-77.

Gotzsch, J. (2017). Getting Creative Again. Awaking Your Inner Creative Self. *The Design Journal*, *20*(sup1), S1072-S1079.

Grimaldi S. (2017). Surprising Longevity. In (Chapman J. ed.) *The Routledge Handbook of Sustainable Product Design*. Routledge/Taylor & Francis.

Gross, A. M. (2016). *Unexpected events and user experience*. (Doctoral dissertation, TU Berlin)

Hatcher, G., Ion, W., Maclachlan, R., Marlow, M., Simpson, B., Wilson, N., & Wodehouse, A. (2018). Using linkography to compare creative methods for group ideation. *Design Studies 58*, 127-152.

Hay, L., McTeague, C., Duffy, A. H., Pidgeon, L. M., Vuletic, T., & Grealy, M. (2017). A systematic review of protocol studies on conceptual design cognition. In *Design Computing and Cognition'16* (pp. 135-153). Springer, Cham.

Ho, A. G., & Siu, K. W. M. G. (2012). Emotion design, emotional design, emotionalize design: A review on their relationships from a new perspective. *The Design Journal*, *15*(1), 9-32.

Hutter, K., & Hoffmann, S. (2014). Surprise, surprise. Ambient media as promotion tool for retailers. *Journal of Retailing*, *90*(1), 93-110.

Kelly, N., & Gero, J. S. (2014). Interpretation in design: modelling how the situation changes during design activity. *Research in Engineering Design*, *25*(2), 109-124.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Ludden, G. D., Schifferstein, H. N., & Hekkert, P. (2008). Surprise as a design strategy. *Design Issues*, *24*(2), 28-38.

Ludden, G. D., Schifferstein, H. N., & Hekkert, P. (2012). Beyond surprise: A longitudinal study on the experience of visual-tactual incongruities in products. *International journal of design*, *6*(1).

Markussen, T. (2009). Bloody robots as emotional design: How emotional structures may change expectations of technology use in hospitals. *International Journal of Design*, *3*(2), 27-39.

Mattelmäki, T., Vaajakallio, K., & Koskinen, I. (2014). What Happened to Empathic Design?. *Design issues*, *30*(1), 67-77.

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, *5*(2), 119-124.

O'Hare, J., Dekoninck, E., Mombeshora, M., Martens, P., Becattini, N., & Boujut, J. F. (2018). Defining requirements for an Augmented Reality system to overcome the challenges of creating and using design representations in co-design sessions. *CoDesign*, in press.

Petersen, R. P., & Buch, A. (2016). Making room in engineering design practices. *Engineering Studies*, *8*(2), 93-115.

Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive‐evolutionary model of surprise: A review of the evidence. *Topics in cognitive science*, *11*(1), 50-74.

Rodríguez Ramírez, E. R. (2014) Industrial design strategies for eliciting surprise. *Design Studies*, Vol. 35, No. 3, pp. 273-297.

Roozenburg, N. F., & Eekels, J. (1995). *Product design: fundamentals and methods* (Vol. 2). Chichester: Wiley.

Roseman, I. J. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition & Emotion*, *10*(3), 241-278.

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In *Appraisal processes in emotion: Theory, methods, research*, 92-120.

Scherer, K. R. (2019). Studying appraisal-driven emotion processes: Taking stock and moving to the future. *Cognition and Emotion*, *33*(1), 31-40.

Scherer, K. R., & Moors, A. (2019). The emotion process: event appraisal and component differentiation. *Annual review of psychology*, *70*, 719-745.

Snider, C. (2014). *Characterising the creative behaviour of designers within the late-stage engineering design process* (Doctoral dissertation, University of Bath).

Srinivasan, V., Song, B., Luo, J., Subburaj, K., Elara, M.R., Blessing, L. and Wood, K., 2018. Does Analogical Distance Affect Performance of Ideation?. *Journal of Mechanical Design*, 140(7), p.071101.

Stompff, G., Smulders, F., & Henze, L. (2016). Surprises are the benefits: reframing in multidisciplinary design teams. *Design Studies*, *47*, 187-214.

Vermaas, P. E. (2009). The flexible meaning of function in engineering. In *DS 58-2: Proceedings of ICED 09, the 17th International Conference on Engineering Design, Vol. 2, Design Theory and Research Methodology, Palo Alto, CA, USA, 24.-27.08. 2009, 113-124*.

Wilson, C. (2013). *Credible checklists and quality questionnaires: A user-centered design method*. Newnes.