CREATIVITY, SURPRISE & DESIGN:
AN INTRODUCTION AND INVESTIGATION

David C. Brown
Computer Science Department, WPI, Worcester, MA, USA

Abstract: This paper weaves together some of the basic ideas about surprise, and creativity, in the context of design. Situations in which design surprise might occur are identified, and a few are discussed in more detail. The goal is to identify areas where additional theoretical, experimental and computational research might be beneficial, leading to the ultimate goal of computational design systems which produce artifacts that are judged to be creative.

Keywords: creativity, design, surprise

1. Introduction

There is no such thing as a creative computational system, only ones that produce artifacts that are evaluated as creative. Hence, to build computational design systems that produce artifacts that will be judged to be creative, we need to understand more about how artifacts get evaluated.

Evaluations have strengths; therefore artifacts may be seen as more, or less, creative. In addition, the evaluation of creativity is multidimensional. Any evaluation in a single dimension results from the evaluator’s biases about how to combine dimensions, where the biases of the evaluator affect which dimensions get weighted more highly. Those biases get changed over time, and change according to the situation. The evaluator may be the designer, someone else (e.g., a user), or a computer program.

Evaluation is based on:

- the knowledge, experiences, context, feelings and preferences of the evaluator;
- the type and degree of exposure to the thing being evaluated (e.g., see, touch, manipulate, or use) (Ludden et al., 2008);
- the “norms” for the relevant population to which the evaluator belongs (e.g., novices, experts).

If computational design systems are going to be reliably judged as creative they will need to be able aim towards creative solutions: hence they should be able to evaluate their own work, both during designing and at completion. Hence evaluation is a key research area.

Many writers have mentioned novelty as one of the dimensions used to evaluate whether an artifact is creative, and some mention value or utility (Boden 1994) (Maher & Fisher 2011) (Sarkar & Chakrabarti 2011). Fewer discuss the role of surprise in evaluation: creative products often seem surprising at first. For example, Besemer (2006) includes it in a set of nine dimensions that she has discovered are used by people to evaluated product creativity (Brown 2008).
As surprise is part of creativity evaluation, computers may need to be able to be surprised by their own work, and be able to judge the likely surprise of others. Note that we are concerned here with cognitive surprise, and not any physical or emotional aspects.

In this paper we describe Boden’s model of some types of creativity, then the eight types of surprise described by Ortony & Partridge. Next we discuss the criterion of radical originality, followed by an exposition of 24 contexts for surprise, finishing with an examination of some of these contexts.

2. Boden’s Model

Boden (1994) writes about a conceptual space: i.e., all the concepts that can result from a synthesis system (human or machine) using its knowledge. The “space” is just the collection of artifact descriptions (e.g., about designs, poetry, or music) that is possible given that knowledge and those synthesis mechanisms.

For convenience, from here on we’ll refer to an “artifact”, with “description” implied, even though for design an artifact is a realization of a description. We will use the term “artifact” instead of “design” in order to try to allow for some generality.

Boden refers to first-time newness (FTN) as a property of those artifacts that have not been generated before (i.e., they’re new), but which fall inside the conceptual space. She considers that these FTN artifacts cannot be judged to be truly creative, as the synthesis system is just “exploring” the conceptual space (i.e., finding new examples from the space, such as a beige milk jug, not a white one). As a consequence, this has been termed exploratory creativity (EC).

Boden (1995) in response to criticism admits that exploratory creativity “can offer surprises comparable to the surprises provided by transformational creativity” (more on the latter below) — note that this does raise the question of why those surprises might be comparable. However, she writes of “mere” exploration to emphasize the difference. Several researchers do not denigrate this kind of creativity quite as much (Ram et al. 1995) (Schank & Foster 1995) (Beghetto & Kaufman 2007) (Brown 2011).

It seems that the perceived weakness associated with exploratory creativity hinges on the fact that FTN artifacts, in principle, either were, or could have been, ‘expected’. The implication is that exploring leads to less surprising results, and, as surprise factors into our evaluation of creativity, those results are less creative.

An expectation is some sort of description of, or proposition about, an artifact (e.g., milk jugs have a handle). An expectation can be satisfied or violated when matched with a description of a newly generated artifact. An expectation violation resulting from such a mismatch produces one of the most common types of surprise. “Mary had a little lamb, wasn’t the doctor surprised” surprises both the doctor, and you, with expectation violations.

So, for FTN artifacts within a conceptual space, we make the assumption (for now) that expectations about those artifacts are similarly limited to that space. Those expectations are deducible in that they could be produced from the knowledge base being used for synthesis. However, Ortony & Partridge (1987) refer to practically deducible expectations, adding a touch of realism, because they are trying to describe a plausible cognitive model.

They do not provide a detailed explanation of what constitutes “practical”, just stating that it should “not require many and complex inferences”. It’s likely that what is practical or not depends on the situation. For example, given a design and plenty of time to examine it visually, and perhaps physically, it is certainly possible to be surprised by inferred expectations about the artifact as more data about it is gathered. As such, “practically” is very much like the term “real time”, in that it depends on the current situation and its temporal constraints. Note that using the term “deducible” is not intended to make any claims about the kinds of reasoning involved.

Clearly, not all expectations about artifacts are necessarily “practically” deducible. Some may be deducible but not practically. As a consequence we can consider two kinds of FTN expectations: practically deducible (PD) expectations and those that are deducible, but not practically deducible (NPD). The amount of effort involved for each is different, as the reasoning chains will be longer,
and/or more complex in the NPD case: i.e., you will get an expectation eventually. The collection of deducible expectations (D) consists of the union of PD and NPD: i.e., \( D = PD \cup NPD \)

The PD expectations can be further subdivided. Imagine working on a design problem, having framed the problem in a certain way. Some of one’s knowledge would be in focus, or active, as it is relevant to the task at hand, while other portions would not currently be in use.

Active knowledge might include recently made decisions, for example. As a consequence of that and other active knowledge, there would be active expectations (A). For example, having decided that a width was 1 mm, there might be an immediate active expectation produced that the length would be less than some amount due to strength requirements. Active expectations are available for matching as soon as an artifact is presented. This is clearly in the PD class, as when matching is done (e.g., with the actual length), no inference is needed: the expectation is already present.

Even if the active knowledge did not produce active expectations, the fact that it is being focused on allows expectation inference to be fast. These active knowledge expectations (AK) are also be in the PD class, as when matching is done and expectations need to be tested, a relatively small amount of reasoning is required to produce them.

One can imagine a class of PD expectations formed from knowledge outside the active knowledge: i.e., not active knowledge expectations (NAK). Deducing those expectations will take a little longer. So, \( PD = A \cup AK \cup NAK \), and each type has the potential for producing useful expectations.

For completeness, it is possible to conjecture knowledge that does not combine with other knowledge and the existing reasoning mechanisms to allow expectations: i.e., expectations are not deducible (ND). That knowledge would not to be useful for synthesizing any FTN artifacts either.

Ram et al. (1995) consider creativity to be in the context of a task and a specific situation. This “pragmatic context” influences and restricts the kinds of the knowledge and the reasoning that are used: i.e., it affects what is deducible. This allows a designer to use strategic (i.e., “meta”) reasoning to change the context, changing what is in D, and hence in PD too.

3. Surprise

There is little detailed discussion about the role of surprise in creativity, the mechanisms that underlie surprise, and surprise in creative design specifically. An important recent review paper about computational surprise is “Artificial Surprise” (Macedo et al. 2009). The seminal paper from AI is by Ortony & Partridge (O&P) (1987), describing the role of expectations in cognitive surprise (N.B., this does not address emotional or physical responses).

O&P introduce eight types of surprise, divided into two categories, one where expectations are practically deducible (PD), and one where this is not possible (NPD): i.e., given the current knowledge something could not easily have been expected. They propose two cases of surprise with PD expectations. An active expectation case (A) where the expectations have been already been produced (i.e., “I expect this”). A passive expectation case allows for situations where expectations can be made active on demand (i.e., “could I have expected that?”). This passive case corresponds to the AK \( \cup \) NAK expectations described above.

O&P suggest that violations of active expectations should produce more surprise than violations of passive expectations, presumably because the knowledge involved in producing active expectations is more relevant to the situation. If active knowledge is the source for active expectations, perhaps that represents the more typical, more frequent propositions.

O&P suggest another distinction that affects the amount of surprise: the typicality of the content of the expectation. For example, an expectation that a desk has a flat top is about something you’d expect of all desks, while an expectation that a desk has drawers is about something typical. So a new artifact can contradict an expectation where the knowledge involved is considered to be immutable (i.e., always true), or an expectation representing something typical. Note that this is a continuous dimension; however, O&P take two values for convenience.
One would expect violations of immutable expectations to be the strongest, and stronger than violations of typical expectations. It is probably also affected by the degree of typicality, the belief in (or evidence for) that degree value, and whether the expectation was active.

There is on-going discussion about how the amount of surprise is produced (Macedo et al. 2009; Maguire et al. 2011). One view favors some form of probability-based surprise, and another considers surprise to be relative to the degree to which “representation updating” is required to make sense of the new event: i.e., if an event is very different from an expectation, but can easily be explained, its surprise will be lowered. In this paper we try to remain as neutral as possible about these views.

In addition to PD expectations, there are also situations where expectations are not practically deducible (NPD). O&P suggest that even the fact that nothing was deduced is in itself surprising.

In the NPD case there can be no ‘active’ expectations. They refer to a kind of passive expectation that can result in “a conflict between the input proposition [i.e., the artifact] and what, after the fact, may be judged to be normal or usual”; i.e., the conflict is “on the basis of deviations from the norms”.

O&P give the NPD example of a finding a person levitating as contradicting “immutable” passive expectations. Having a rock come through one’s window is given as an example of contradicting “typical” passive expectations: but it too (under normal circumstances) is not practically deducible.

Notice that the conflict they describe is between the input (e.g., a person) and a proposition about what is “normal or usual”. However, that judgment requires some kind of inference (e.g., induction from examples). But if it is inferred then this makes it practically deducible! That proposition about what is normal is a passive expectation, and must be in the PD class. This contradicts O&P’s claim that this kind of surprise occurs in the NPD situation. Hence we must conclude that contradicting norms is a special case of PD, and that given O&P’s model there are no NPD examples of surprise. This seems wrong initially, but clearly what is practical can be situation dependent: i.e., there are examples, but not in this situation.

Interestingly, despite previously following O&P very closely, Macedo has pared down his model of surprise (Macedo et al. 2009) to remove NPD cases, without stating it as explicitly as we have above.

4. Expectations about Surprise

In Macedo & Cardoso’s (2001) multi-agent system, agents produce descriptions of new items for other agents to inspect. Each agent considers a variety of possible new items, evaluating each for the amount of surprise it is expected to cause in other agents. As there is potential benefit from causing surprise, the most surprising design is chosen.

Normally, this evaluation would require a model of the other agents’ knowledge: i.e., you have to know the other agents’ norms in order to violate them. In that system it can be done because there is a single, shared, frequency-based scheme (i.e., the norm) for calculating the amount of surprise, and therefore no need for models of what other agents know.

Designer agents form expectations for each possible new item about how much other agents will be surprised. We will refer to these expectations as surprise expectations (S). We conjecture that a human or computational system making design decisions can be motivated by the goal of producing creative artifacts, hence they might intend to surprise, and could have S expectations.

The S expectations appear to be ‘active’. They are not deducible at the design knowledge level, but should be at the design process meta-level. S expectations will not be included with type A expectations, as despite being active they are not directly about the design.

In Macedo & Cardoso’s system, because of the shared model of surprise, the S expectations will be confirmed. However, in general, if the evaluator is not surprised by the design, then the designer’s S expectation fails, causing the designer to be surprised (“Wow! I thought they’d be surprised!”).

5. Radical Originality

Boden’s view of the newness needed for what she considers to be true creativity, and not “mere” exploration (EC), is the kind of “radical originality” that results from changing the conceptual space. As the space needs to be transformed, this is referred to as transformational creativity (TC).
The resulting new, creative artifacts are in the transformed space, but not in the original space. They could not have been expected using the original knowledge (for the old conceptual space), because those expectations do not apply to the new conceptual space by definition: i.e., expectations about complete new artifacts are not deducible for the new space using the original knowledge.

Boden suggests that these new artifacts, resulting from the transformation, would be “surprising". She does not say what this means, nor for whom this will be a surprise. We assume that she means that expectations are violated, and that this is at least surprising for the external evaluator.

The transformation might be produced by relaxing constraints, abduction, induction, analogy, or by some other method. Other possibilities include mental simulation, reinterpretation, emergence, explanation, and merging/combining knowledge from other less pragmatically related areas. For more discussion of transformations see Ritchie (2005; 2006) and Wiggins (2006).

Boden also writes about combinatorial creativity. While EC and TC cases are defined in terms of the space of possible concepts, combinatorial creativity is a way of generating new concepts by combining existing ones. Note that this can transform the conceptual space. As this type of creativity is more about ‘process’ it is categorically different from EC and TC, and appears to be an example of how to produce a new artifact in a transformed space.

Consider the transformational creativity of a designer who has transformed the conceptual space: she now has new, recently formed knowledge. As before, we assume that the designer can now form an expectation about any new design that they might generate using this new knowledge.

Ritchie (2005) is unusual in that he uses “the perspective of the individual assessing the artifact”, thus putting the onus for transformation on the evaluator: i.e., the evaluator needs to transform his or her own conceptual space so that the designer’s artifact ‘fits’ into the transformed space. This resonates with Schank’s theories of “creative explanation”, as well as the theory (Maguire et al. 2011) that surprise is concerned with revising knowledge to explain the difference between the expectation and the artifact.

**Table 1. When and How Surprise Occurs**

<table>
<thead>
<tr>
<th></th>
<th>EXPLORATORY</th>
<th>TRANSFORMATIONAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DESIGNER</td>
<td>EVALUATOR</td>
</tr>
<tr>
<td>PARTIAL DESIGN</td>
<td>A 1 4 7 10</td>
<td>AK 2 5 8 11</td>
</tr>
<tr>
<td></td>
<td>NAK 3 6 9 12</td>
<td></td>
</tr>
<tr>
<td>COMPLETE DESIGN</td>
<td>A 13 16 19 22</td>
<td>AK 14 17 29 23</td>
</tr>
</tbody>
</table>

**6. Contexts for Surprise**

At this point we can ask in what situations surprise might occur, and whether the classes of expectations, A, AK, and NAK, can affect the type of surprise that might be produced if each were violated. A key issue is how surprise can occur for EC, in addition to TC. The other variables that affect the context for surprise are whether the expectation is used with a complete, finished design or with a partial design during the design process. As the designer may or may not be the evaluator, both cases need to be considered: i.e., who gets surprised. This produces 24 contexts for surprise (Table 1).
For example, the surprise in cases 1, 4, 7, and 10 represent situations where active expectations about a partial design are violated, while 2, 3, 5, 6, 8, 9, 11, and 12 are passive expectation violations. It is most likely that only the designer will see partial designs, while both the designer and the evaluator see complete designs. The most significant outstanding question is “how can a designer be surprised by their own design?”.

7. Discussion

Ritchie (2005) discusses what Boden might mean by a conceptual space, distinguishing between an artifact-set, and a space-definition. The former is the actual set of artifacts that constitute the space, while the latter is “a more compact definition of possible artifacts”. The space-definition is knowledge that can be used to generate new artifacts, while known artifacts in the artifact-set can be used for case-based reasoning as well as expectation formation.

Imagine that space-definition knowledge can be modeled in rule form. Clearly, a designer can know the rules, but not necessarily know in advance all the results of successive applications of these rules. It is highly unlikely that much of the space is already known: probably occurring only in very routine situations (Brown 1996). As Boden (1995) points out, exploration can take us to “regions that were previously unsuspected” and show us boundaries “in surprising places”: i.e., the designer is surprised.

Using natural language as an example, the phrase “half a pair of purple trousers isn’t really better than none” might be in the space of sentences (i.e., artifacts), but it would fall into a “previously unsuspected” region: before now, that is. This corresponds to exploratory creativity.

Another possible argument for why a designer might surprise themselves, in both EC and TC cases, can be made if we consider a conceptual space to be formed from concepts. Rosch (1999) describes the “Graded Structure/Prototype” view of concepts as: not having clear cut boundaries; not having necessary attributes to determine membership; having gradations of membership; containing a rich amount of information about some situations; and being very flexible. She considers concepts to be “situation based and participatory rather than identification functions”, strongly tied to contexts and situations. This makes the conceptual space much fuzzier and unpredictable than, for example, something generated by a formal system. Given this view it is much easier to envision that a designer might be surprised by some intermediate or resulting design.

One possible situation where a designer might be surprised by their own partial designs is if they are sketching. Sketching allows the possibility of recognizing unexpected associations with other design concepts, or finding emergent properties (Suwa et al. 1999). In this way their expectations (probably “active”) will be violated resulting in surprise about a partial design. It is even possible that such associations might be ‘fuel’ for a transformation of the conceptual space.

A general issue for the TC case is whether expectations can be formed as easily for the transformed space as for the original conceptual space: after all, the original space is more familiar and more explored already. It seems likely that until the transformed space becomes familiar, TC artifacts will cause surprise for the designer as she will tend to produce expectations from the original conceptual space. This is clearly just a hypothesis, but a plausible and interesting one.

Referring back to the natural language example, the phrase “lovely chubby babies sleep quietly” could be recognized as the result of known rules (EC), and even expected, but in the TC case, “awkward blue concepts feel furiously” is surprising, even though we can infer that relaxation of semantic constraints must have been involved in the transformation of the space.

Let us assume that an external evaluator shares the same original knowledge, and the same original conceptual space, as the designer. They will both be able to produce active or passive expectations about any design in the original conceptual space: although not necessarily the same at the same time. However, using that original knowledge, they will not be able to produce expectations about complete designs in the transformed conceptual space. However, despite that, it is possible that some expectations about properties of designs in the transformed space are shared with those in the old space, as not everything about the new designs will be radically new. As Boden (1995) points out, “some aspects of the previous space need to be retained”.
Consequently, the designer, using their original untransformed knowledge, should be able to try to evaluate their own new TC designs for creativity. Some of their expectations will be satisfied, but some will be violated due to the changes brought about by the transformation. Their surprise will result in the same way as an external evaluator’s might if they share the same original knowledge.

Alternatively, the evaluator might have enough design ability and experience to have an expanded or different conceptual space, able to produce active or passive practically deducible expectations about some TC design solutions. So, depending on the overlap between the designer’s transformed space and the evaluator’s expanded space, in some situations the evaluator would not be surprised by TC solutions, and sometimes they would have surprise from expectation failure.

In general, however, the designer’s knowledge may be quite different from the evaluators (e.g., customer, professor, CEO, etc.). For example, in case 22 (Table 1) a complete TC design is evaluated by an evaluator with an active expectation and she is surprised. Active expectations can be caused by prior knowledge of that product type, with knowledge of examples, probably plus extensive knowledge of typical similar products. The correctness of an expectation depends on the closeness of the design to existing products. Note that if the evaluator has access to intermediate versions of the design (i.e., partial) then this might allow them to produce additional expectations.

The evaluator could even know requirements and have mentally produced a rough design themselves, producing expectations about many of the design decisions: much as was done by the Active Design Documents system (Garcia et al. 1993).

In situations where redesign is being done — perhaps due to the evaluator having done an evaluation of which aspects of a rough/conceptual design might be seen as creative, as Besemer (2006) suggests — they might have quite strong, active expectations about the outcome.

When the evaluator has the same original knowledge as the designer then their expectations will be mostly limited to the untransformed design space. To be surprised the evaluator must be faced with properties of the TC design that do not arise from old knowledge. In situations where the designer has more knowledge than the evaluator, even exploratory creativity might be quite surprising to the evaluator, and TC designing is sure to be. In a situation where the evaluator has more knowledge than the designer, it is more likely that the evaluator’s expectations will be satisfied (i.e., they’ve seen it before). So while it might be P-creative (Boden 1994) for the designer, it will still not be surprising for the evaluator (similarly for cases 23 & 24, and passive expectations).

Maguire et al. (2011) emphasize the difference that this kind of sophistication can make; pointing out that a “flying rabbit” would be very surprising for an adult, with strong (immutable) expectations, but “more easily accepted” by a child. However, they propose that adults tend towards generalized representations that allow easy integration of surprising events. Children however find most things surprising, as what representations they have are still quite specific and hard to integrate with. An interesting issue is whether the expectations themselves can vary from general to specific, thus changing the matching process with the artifact, perhaps changing the way expectations are “violated”.

Horn & Salvendy (2006) confirm that expectations influence product creativity evaluation. If a product falls short of expectations it can lead to “disappointment” and “dissatisfaction” but pleasure and emotional arousal affect judgments of creativity. These effects are claimed to apply across each dimension of the evaluation (e.g., novelty).

8. Conclusion
This paper has attempted to present and weave together some of the basic ideas about surprise, and creativity, in the context of design. Situations in which design surprise might occur have been identified, and a few were discussed in more detail. This provides a framework in which to consider situations where designers or evaluators might be surprised.

The background goal of this work is to identify areas where additional theoretical, experimental and computational research might be beneficial, leading to the ultimate goal of computational design systems that produce creative artifacts.
References


