GUIDING COMPUTATIONAL DESIGN CREATIVITY RESEARCH

DAVID C BROWN
Computer Science Department, WPI, Worcester, MA 01609, USA

Abstract. As the existence of creativity is a judgment, relative to personal or group norms it is helpful to understand what factors inform that judgment. Creativity can be assessed using either the designed product or the design process. Judging the process is much more problematic than judging the product. Besemer's Creative Product Semantic Scale for product assessment provides three dimensions, with nine attributes that can be used for judgments. As expected, "novelty" is one of the dimensions, but its three attributes provide more subtlety. This paper discusses how creative product assessment scales might provide important clues about how computational systems can be built or controlled in order to produce creative products.

1. Introduction

This paper is concerned with how Computer Science (CS) can assist with "studying design creativity". While creative design is clearly part of much CS activity, such as system, algorithm and program design, it is rarely taught from a creativity point of view, or studied. However, it is clear that CS techniques could be used to collect, analyze and visualize data about creative design in CS or other areas. CS can also aid the study of design creativity by designing and building support tools that, for example, support collaboration, manage complexity, keep history and rationale, and support exploratory search (Shneiderman 2007; Lubert 2005).

An alternative approach involving CS is to try to understand design creativity by "mimicking" it using the computer. That approach is the focus of this paper. As design is a human activity requiring intelligence its computational study falls within the realm of Artificial Intelligence (AI). An underlying assumption in AI is that making hypotheses about knowledge and reasoning, and then building systems that embody them, is a good way to investigate intelligent activities. We want these computational models to provide reasonable explanations of design reasoning, and to be consistent with observations. Such models may provide explanations when no others at the same level of detail exist.

Much of the research in the "AI in Design" area proceeds in this way: for example, modeling routine design, analogical reasoning, or configuration, and then building systems that instantiate those models.

These systems can be used to experiment with the consequences of changing the knowledge or the problem-solving used. The resulting changes in the designs (and design reasoning) can be evaluated by experts or compared with existing designs to attempt to "validate" the model.

As creativity is an intelligent activity, AI-based studies of creativity, often called "Computational Creativity", proceed in much the same way.

Designing in general is a complex, opaque activity, and the same is true of creativity. Producing hypotheses about either alone is hard: producing hypotheses about both together is even harder. Computational Design Creativity (CDC) will need to be studied incrementally, with many hypotheses investigated.

It is extremely important to note that creative designing is, after all, a form of *designing*, and therefore any computational model of 'creative' designing must be firmly based on a strong computational model of 'ordinary' designing. The alternative is that creative designing is distinctly different from designing. However, neither I, nor most of the authors cited below, believe that to be true.

Creativity in design can be detected and assessed by looking at either the designed product or, when possible, the design process. As the existence of creativity is a judgment, relative to personal or group norms (Boden 1994; Amabile 1983), it is beneficial to understand what factors inform that judgment.

The first place to obtain clues as to what factors might be involved is from definitions of, and discussions about, creativity. The second place is from existing research in Computational Creativity: unfortunately, space does not allow us to review that in this paper.

The third place to look is at research that directly concerns how products are judged to be creative. Fortunately, the factors used to make such judgments have been studied. The Creative Product Analysis Model (CPAM), with its associated Creative Product Semantic Scale (CPSS) for product assessment (Besemer and Treffinger 1981; O'Quin and Besemer 1989; Besemer 2006), provides three dimensions, with nine attributes that can be used for judgments. As expected, "novelty" is one of the dimensions, but its two attributes provide additional subtlety.

A hypothesis of this paper is that the scales used by research in creative product assessment can provide important clues about how CDC systems might be designed in order to produce creative products.

The main body of this paper continues by investigating the consequences for CDC systems of several selected pieces of work about creativity. With that foundation established, it moves to an introduction of the CPAM, describing the main factors and their characteristics. The many consequences of that model for CDC systems are described next. The paper then briefly mentions some other existing measures that might be used for CDC systems, before proceeding to a summary and a concluding discussion.

2. The Creativity Literature

In this section we will review a very small selection of the huge body of work about creativity, attempting to highlight aspects that might be useful for the construction of CDC systems that could be used to study creativity. The work is drawn from a variety of disciplines.

2.1. AMABILE (1983)

Amabile provides the following definition of creativity:

> "A product or response is creative to the extent that appropriate observers independently agree it is creative. Appropriate observers are those familiar with the domain in which the product was created or the response articulated. Thus creativity can be regarded as the quality of products or responses judged to be creative by appropriate observers, and it can also be regarded as the process by which something so judged is produced."

She also argues for "degrees of creativity", from "garden variety" (use of familiar algorithms) to "historically significant" (use of heuristics and new "cognitive pathways"), instead of it being a dichotomy. She considers that to be judged as creative by an observer a product must be novel, appropriate, useful, correct or valuable. In addition she argues that the design task should be heuristic and not algorithmic: i.e., the path to the solution is not "clear and readily identifiable".

Amabile stresses that ingredients in creativity are: education; strong, appropriate cognitive skills; and "playful" freedom from constraints. She lists the following as creativity-relevant skills: breaking away from functional fixedness; abandoning old, unsuccessful problem-solving strategies; keeping options open as long as possible; suspending judgment; seeing relationships between diverse bits of information; accurately remembering large amounts of information; breaking out of well-used scripts, and actively examining them; having heuristic knowledge that promotes creativity; and productive forgetting.

Amabile argues that having more domain knowledge allows "more response possibilities" and more ability to judge those alternatives. She states that "although it is possible to have too many algorithms, it is not possible to have too much knowledge". However, she makes it clear that the knowledge should be organized according to “general principles” in order to increase its utility.

2.1.1. Additional Comments about Amabile’s Ideas

Amabile lays the groundwork for CDC systems by arguing for a variety of ingredients, most of which suggest computational approaches. Most of these ingredients will be discussed below, but one key aspect to her list is the need for several types of meta-level activity, such as ensuring least commitment, detecting functional fixedness, and observing plan/script progress. Also important is her stress on having lots of well-organized knowledge.

2.2. BODEN (1994)

Boden argues that creative ideas must be *new* and *valuable* to be interesting. However, she wants to “distinguish radical novelties from mere ‘first-time’ newness”, as the latter can be generated by a system (perhaps using rules) that “underlies the domain and defines a certain range of possibilities”. This “conceptual space” defines what *could be* produced by a system, resulting in newness that is in some sense expected. However, the conceptual space can be changed by transformations in order to allow radical originality. She provides the actions of dropping or inverting constraints as examples of such transformations, and suggests that some “meta-representation” of such constraints would allow their modification. It’s worth noting that Wiggins (2003) and Gero (1994) provide more detailed accounts of what transformational creativity might be.

Boden suggests that an understanding of how “novel combinations of old ideas” come about, coupled with a theory of analogical reasoning, would provide a substantial start to a theory of creative reasoning.

The other major aspect of Boden’s paper is the distinction between two senses of creativity. Psychological creativity, or P-creativity, occurs “if the person in whose mind it arises could not have had it before”. By “could not” she means that producing the valuable idea required a transformation to their conceptual space: i.e., P-creativity is personal. Historical creativity, or H-creativity, requires that “no one else, in all human history, has ever done it before”. H-creative ideas are necessarily P-creative too.

Boden suggests that different individuals or groups may vary in their assessment of the value of an idea. In addition, they may vary in their assessment of newness. The consequence is that the judgment of creativity by others (including systems) can also vary. A very important related point she makes is that recognition of creativity by a computer system is necessary if that system is to appear to be creative.

2.2.1. Additional Comments about Boden's Ideas

The notion of transforming a conceptual space, by constraint manipulation for example, to extend what "could be" generated is a credible computational device. Like Amabile, she too suggests meta-level reasoning (about constraints) as an ingredient of creative reasoning.

Boden's definition of H-creative, as stated, appears impractical for products. The judgment of whether *anyone* else "has ever done it before" implies comparison to a huge body of knowledge (in people or books). The variables for that knowledge are the length of time the record represents, the number of sources involved, and the level of expertise represented. For a CDC system it is good to have high values for all of these variables.

We can assume that "experts" know the most about the historical record and about major historical, creative changes in a product class. Therefore a large number of experts could judge whether something is H-creative. That knowledge could be made available for CDC system use.

What if you only had a few experts? Or what if the people weren't experts? Would they then be able to judge creativity?

One option is that H-creativity might be seen as relative to the size of the group making the decision about originality. For example, the world's university professors might observe H-creativity relative to the world; all the members of a professional academic society might observe H-creativity relative to the society; or all the teachers in a school may observe H-creativity relative to the school. Clearly the selection of the group providing a CDC's knowledge will be important, as it defines the 'scope' of the judgment.

It is reasonable to argue that individual H-creativity is also a useful concept (i.e., a group of one), despite Boden's objection to "mere 'first-time' newness". In fact, judging one's own or another's P-creativity might be so hard that individual H-creativity is all you have.

2.3. EYSENCK (1994)

Eysenck proposes that:

> "Creativity denotes a person's capacity to produce new or original ideas, insights, inventions, or artistic products, which are accepted by experts as being of scientific, aesthetic, social, or technical value."

In addition he requires it to have "acceptability or appropriateness".

Eysenck distinguishes between "private novelty" (new to the individual generating the item) and "public novelty" (new to "everyone"), as well as "creativity as a trait" (characteristics of a person that lead them towards private novelty) versus "creativity as shown by productivity" (i.e., the production of works of public novelty). He also distinguishes between creative products and creative processes, as well as creative situations.

Regarding the trait, he argues that there is evidence to support a relationship between creativity and individuals who demonstrate “loosening of associative thinking”, “a quality of over inclusiveness”, and “a failure of inhibition that allows less relevant thoughts to intrude into the problem solving process”. This allows “ideas and associations” to “become relevant that would not appear to be so for the ordinary person”. However, creativity requires these additional associations to be narrowed by “critical assessment” to just the most promising.

2.3.1. Additional Comments about Eysenck’s Ideas

Eysenck’s private/public distinction is very similar to Boden’s, with similar consequences. His creative process/product distinction is also important, as it raises the question of how the two are related, as well as the question of whether the same assessment criteria can apply to both process and product.

Clearly, creative processes can lead to creative products. Using Eysenck’s definitions, it *is* possible for not very creative processes to lead to creative products from the point of view of private novelty. Whether “experts” would judge them as having value, or as showing public creativity, is dubious, but not impossible. Conversely, highly creative processes might lead to previously obtained public results, and perhaps even previously obtained private results. The consequences for computational systems appear to be that while creative processes are desirable, they are not necessary to produce “creativity as shown by productivity”.

With regard to “creative situations”, Eysenck was referring to social context, and the influences on the individual that might affect their performance. For a computational system, this suggests either a multi-agent system that directly models those multiple influences, or some way of influencing the system so as to activate portions of the system’s knowledge before or during creativity so that they get priority when the system is reasoning associatively.

The qualities of divergent thinking that Eysenck identifies suggest that computational systems need to ‘ask the right questions’ as they proceed, so that they allow and encourage the reasoning skills that tend to lead to creativity to be utilized. Such questions would be the sub-problems and sub-goals that the system produces. Talbot (1997) quotes Ekvall (1995) as writing “unclear goals were contributing to the climate that made radical innovation possible”: i.e., lack of precision, and/or more abstract descriptions, can be beneficial. In addition, the associative reasoning used may need to be looser than it could be in order to take advantage of such goals.

An important point Eysenck makes is the need for “critical assessment”. Hence, a CDC system needs such abilities. He points out that most creative people are usually highly focused in one technical or artistic area, displaying public creativity requiring “special knowledge that may take years to acquire”. A system that can prune distracting well-known or weaker options will require a lot of specialized knowledge. In a more extreme version of this argument, Weisberg (1999) argues that “we do not need special theories to explain creative thinking” as the key issue is “determining the knowledge that the creative thinker brings to the situation”. However, many writers point out the tension between having knowledge that allows good decisions, and being constrained by it so much as to reduce creative reasoning.

2.4. DASGUPTA (1994 & 1996)

Dasgupta defines a design as “psychologically novel” (PN) if the designing agent “believes that there exists no other [design] in his personal knowledge body that is identical”. If the agent believes that the design “adds significantly” to the relevant community’s public body of knowledge then it is also “psychologically original” (PO). However, if the *community* believes that the design is new then it is “historically novel” (HN), while if they believe that it adds significantly then it is “historically original” (HO). This leads to different levels of creativity: PN-creative, PO-creative, HN-creative, and HO-creative.

Dasgupta claims that associative retrieval of “related” knowledge, and the willing ability to “roam freely across the span of knowledge”, is vital for creativity. Like many others, he claims that the “combination or association of apparently unrelated ideas” (bisociations) are an ingredient of creativity. However, there is need to identify “misfits” and use “elimination heuristics”.

Dasgupta believes that creative designing is knowledge-intensive, opportunistic, and goal-directed, with goals that are “provisional or tentative”. However, he feels that creative design reasoning processes are much like any other reasoning processes. Problem recognition and product inadequacies lead to problem formulation: i.e., new goals. Searches are carried out for appropriate knowledge and for solutions in a problem space. The knowledge used may be “compiled” or not.

Technological knowledge, which is domain knowledge plus operational principles, contributes to the knowledge intensive characteristic. An operational principle is “any proposition, rule, procedure, or conceptual frame of reference about artifactual properties or characteristics that facilitate action for the creation, manipulation, and modification of artifactual forms and their implementation”. Operational principles drive creativity, and include heuristics and expectations.

Knowledge also includes records of errors made and the failures that resulted from them. Dasgupta cites Petroski’s four error types: “flawed conceptual design” due to inadequate criticism, and weak evaluation/analysis; “oversimplified assumption”; “misunderstood scaling effects”; and “tunnel vision”, i.e., focusing on one aspect of design.

2.4.1. Additional Comments about Dasgupta’s Ideas

Note that Dasgupta separates out the factor of who decides on the creativity (agent, or group), from the contribution that the design makes, with mere novelty at one end of the scale, with novel but a significant contribution at the other. Clearly, PO-creative and HO-creative are the ones that count. He notes that these creativity decisions are all judgments, and all made after the fact. It is still the case that a set of decisions were made during designing that led to the design being PO or HO. It seems likely that those decisions were affected by the designer’s personal body of knowledge as well as his or her understanding of the state of the community’s body of knowledge. One key issue is how to judge what “adds significantly”, and whether this can even be judged during designing.

While Dasgupta considers that any “thought product” may be classified as creative in this way, for design, the final design description is what counts: e.g., a designer may produce HO knowledge (a thought product) on the way to producing a design that is not HO.

If the design activity is to be exploratory, then ideas generated are going to be provisional. This requires some way of describing the nature and sources of uncertainty about an idea.

2.5. WARD & SMITH (1995, 1997 & 1999)

Ward et al. (1997) focus on “‘normative’ creativity”, the “generative potential that is inherent in … most normal human brains”, but acknowledge “the possibility that the mundane and the exotic [forms of creativity] represent endpoints on a continuum of human creativity”. They believe that the same sets of mental processes underlie both.

They list some conceptual processes that they feel are “crucial to creative thought”. These include Conceptual Combination, Conceptual Expansion, Analogy and Mental Models, and Knowledge.

They argue that combination of concepts to produce a novel concept involves “comparison, construction and integration processes”, and that the failure to “find a satisfactory coherent interpretation” might actually lead to a search that results in creativity. They also note that recognizing other kinds of mental blocks, and handling them by approaches such re-representation, is a key principle (Smith et al. 1995).

Expansion occurs when people “construct, stretch, extend, modify and refine” concepts. However, expansion tends to be “driven by the characteristic properties of known concepts”, especially recently encountered instances.

Analogy is often suggested as an ingredient of creative reasoning. They distinguish between infrequently occurring, but highly creative, “far analogies” across widely separate domains, and “near analogies” within a domain or between highly similar systems. Mental models, as well as visualization and creative imagery (Ward et al. 1999), allow complex reasoning and can support discovery.

Ward & Smith argue that creativity is:

> “…an outcome of subsets of those and other processes acting in concert to expand the frontiers of knowledge and conceptualization in a given domain”

and that the outcomes are “rooted in existing knowledge”. They stress that creativity requires “cognitive labor”, in order to fully explore mappings and relationships in order to tease out sensible interpretations and interesting mappings. This might be especially difficult in the face of the “cognitive inertia” of existing knowledge structures to resist change.

2.5.1. Additional Comments about Ward & Smith’s Ideas

The key consequence for computational systems is that creativity results from a variety of different sets of conceptual processes affected by context, and that a “single overarching theory of creativity” isn’t very likely. Smith (1995) suggests that some creative thinking might be guided by plans that are retrieved or constructed. In addition, retrieving appropriate knowledge might be guided by being in the context of a particular plan. Plans represent subgoals, and also contain other subgoals.

2.6. MINSKY (2006)

Minsky starts his discussion of creativity by quoting Linus Pauling’s “The best way to have a good idea is to have lots of ideas”. Clearly this is a goal for a CDC system. He suggests that there is “no single, unusual trait” that accounts for high creativity, but instead that it is due to the unusual combination of strength in common traits. These include considerable expertise, persistence, a large variety of ways to think, more novel thinking, resistance to irrelevant goals, rejection of “popular myths and beliefs”, spending less time on unproductive ideas, and learning more from less experience. Minsky also suggests that good “mental management” is important: i.e., they “organize and apply what they learn”.

Minsky argues that creativity may stem from enjoying the “discomfort” of not following the ‘best’ or best known method for a goal, and that this leads to “exploration pleasure”. Interestingly, Amabile has concluded that positive feelings are strongly related to creativity.

He argues that creative thinkers generate a lot of new, and perhaps novel ideas, they are distinguished by “how effectively they can select which new ideas to further develop", and that they tend to suppress those with "too much novelty". They also need to determine the constraints that indicate "plausible things to try". Minsky, as well as Besemer, mentions "elegance" as a sign of a possible successful candidate, and that evaluation can confirm this, or indicate where it might need repair.

He considers that "negative expertise" (i.e., what *not* to do) is a very large part of people's knowledge, based mostly on learning from mistakes: i.e., we do positive things by not doing negative ones. Critics are resources that can "recognize some particular kind of potential mistake", either during the process (Corrector), while planning it (Suppressor), or even before it can reach the planning stage (Censor). When being creative, some critics need to be "switched off" so that fewer hypotheses will be rejected. Critics can also "recognize successes and promising opportunities": these are known as Encouragers.

Critics can also recognize particular mental conditions, and then activate appropriate Selectors that "activate a way to think". The "ways to think" include reasoning by analogy, divide and conquer, re-representation, planning, simplifying, and changing the level of abstraction, etcetera. Minsky also argues for the need for mappings between "problem types", such as a failing prediction or lack of knowledge, and the appropriate ways to think.

Minsky argues for the necessity for a rich, interconnected variety of different types of knowledge representation with multiple versions of the same concept at different levels. Different representations ought to be better for different kinds of reasoning about different problems. Each piece of knowledge should be associated with the goals it might help with, the situations for which it might be relevant, what subgoals must be achieved before use, what it has been used for before, what harm or good has its use produced, the cost of use, what methods it works well with, and its common exceptions or bugs.

Differences play a big role in reasoning. Their detection helps us to try to achieve goals and to interpret information (where differences are changes in lower level representations), while explicitly representing similarities and differences allows us to break out of rigid hierarchical representations.

Minsky argues that reasoning by analogy is "our most usual way to deal with problems", as new situations are never quite the same as old ones. Differences play a key role in reuse of old methods. He points out that knowing which differences to ignore affects how similar things appear to be, but that the importance of each difference depends on current intentions and goals. Hence analogies depend on them as well.

Many other authors have discussed the role of analogy in design and in creativity: see (Goel, 1997) for an introduction. Several authors have noted that analogical reasoning can be done using representations of structure (S), behavior (B) or function (F) (e.g., Balazs & Brown 2001). Engineering and architectural design theorists have separated S, B and F in their (sometimes prescriptive) models of designing for a long time. Since about 1985 AI researchers have used SBF (or FBS) computational models of devices (Erden et al. 2008). Gero and Kannengiesser (2007), using a situated refinement of an earlier SBF-based model of designing, specify 20 activities that can occur between and within S, B or F representations, and which may lead to creativity. They note that some of the activities can be done by analogical reasoning.

3. Creative Product Analysis Model

Despite Boden's claim (1994) that "In general, one cannot assess creative ideas by a scalar metric", the Creative Product Analysis Model (CPAM) aims to do something similar (Besemer & Treffinger 1981). It is the basis for a well-validated product creativity assessment instrument called CPSS (O'Quin & Besemer 1989) (Horn & Salvendy 2006) that is demonstrating its utility in current, practical use. The model has three main dimensions (also known as factors): *Novelty*, *Resolution* and *Style*. Each of these factors has between 2-4 characteristics that further refine them.

Novelty is "the extent of newness in the product" and refers to the "number and extent" of the new processes, new techniques and new concepts included in the product. It also refers to "the newness of the product both in and out of the field" (Besemer 2006). This is the factor that everyone includes when the topic is creativity.

Resolution is "the degree to which the product fits or meets the needs of the problematic situation". This factor enforces the fact that, for products at least, new but bizarre objects aren't seen as creative, as products are usually associated with an intended function, and therefore being 'useful' is prized. When viewed as art, Salvador Dali's celebrated, surrealist Lobster Telephone is creative, but when viewed as a product it has limited utility.

Style is "the degree to which the product combines unlike elements into a refined, developed, coherent whole, statement or unit". Besemer (2006) refers to this factor as "how the product presents itself", the "product's personality". It affects how creative the product is perceived to be, and may even impact how novel it seems. For example, a telephone *covered* in lobster scales would create a very different impression than a phone covered in leather.

Novelty:

Surprising: “The product presents unexpected or unanticipated information to the user, listener, or viewer”. This is the aspect of the product that catches the user’s attention.

Original: “The product is unusual or infrequently seen in the universe of products made by people with similar experience and training”.

Resolution:

Logical: “The product or solution follows the acceptable and understood rules for the discipline”.

Useful: “The product has clear practical applications”.

Valuable: “The product is judged worthy because it fills a financial, physical, social, or psychological need”.

Understandable: “The product is presented in a communicative, self-disclosing way, which is ‘user-friendly’”.

Style:

Organic: “The product has a sense of wholeness or completeness about it. All the parts work well together”.

Well-Crafted: “The product has been worked and reworked with care to develop it to its highest possible level for this point in time”.

Elegant: “The product shows a solution that is expressed in a refined, understated way”.

4. Consequences of the CPAM

For now, let us imagine that we can devise ways to use the various factors included in the CPAM as goals used in a CDC system (CPAM factor goals). By driving the system towards these goals, designs that are evaluated as more creative should be generated. An obvious consequence of trying to use CPAM in this way is that there are three main dimensions of influence, and a large space of measures in which a design might sit. The system would need to know which parts of the space are the most productive for generating creative products.

What CPAM can provide is a product profile with scores for each of the nine characteristics (Besemer 2006). The shape of the profile histogram indicates variations in creativity. Products can be ‘aimed’ towards a goal profile shape, depending on which characteristics are more valued (Besemer 2008). Consequently a computational design system might be given a goal profile to influence its reasoning, or have one built in. For example, as extreme novelty brings doubt in evaluators’ minds, leading to low scores for other characteristics as a reaction, extreme novelty might be avoided by lowering that part of the goal profile.

Besemer (2006), while giving an example of product improvement, generates alternative ideas for the product, then suggests considering the "weakest CPAM facet" of each of the ideas. The ideas can then be improved by tweaking them to try to address the weaknesses. This is clearly an example of the sort of activity we are proposing for a CDC system.

One proposal is to target particular places in the CDC where actions will affect portions of the resulting creativity profile. This requires an analysis of what 'causes' could produce certain positive 'effects' in the profile. In this approach the resulting profile can be produced by a human judging the CDC's output. However, we hypothesize that it should also be possible for the CDC to form the profile. This allows the CDC to evaluate its own creativity. It may also be possible to use the profile to provide evaluations of partial designs, thus shaping the CDC's design process.

Boden (1994, p.112) argues that creativity "cannot be measured". However, the Creative Product Analysis Model doesn't merely rate products on a linear scale from uncreative to creative, but uses many dimensions. The advantage of any sort of metric is that the values do not need to be "correct", just as long as it provides relative consistency, allowing reliable comparisons to be made between products in the same general category.

In addition, Boden seems to be focused more on artistic creativity, where requirements may not exist, and constraints on the artistic "product" are usually looser or non-existent. As already suggested (i.e., the Lobster Telephone), evaluating an object in the context of Art is quite different from the context of Product. As creativity is often associated with the arts, it isn't surprising to find that many computational creativity systems produce artistic artifacts, such as pictures or music. Such results are evaluated by artistic standards and by the perceiver's "taste". This means that there is less chance of a generated artifact being judged in an absolute fashion, so the standards are softer and therefore easier to satisfy (Brown 2007). As Boden (1994) quotes, "I don't know anything about art, but I know what I like".

Hence, developing computational systems for and writing about creativity in the arts and products contexts ought to be distinguished, as Christiaans (1992) also concludes. As CPAM was developed and verified for products only, it isn't clear how valid its application would be for other contexts.

This raises the issue of whether CPAM can be used to assess the creativity of design processes. Besemer (2008) feels that it should be able to, however characteristics such as "understandable", "well-crafted" and "elegant" seem very product-dependent, even though there appear to be analogous measures for processes. In the rest of this paper we will only consider the evaluation of partial or complete designs, not processes.

A major issue is whether the impact of every design decision on every CPAM characteristic can be determined, or even whether it needs to be assessed at every decision point. Where expectations exist (Grecu & Brown 1998), or expectation failures have occurred (Brown et al. 1992), there may be some appropriate knowledge. In other cases heuristic relationships may exist. Even if measurements of each characteristic are not possible during designing, it should be possible to positively influence the scores if those causal relationships are known (even partially). A confounding issue is whether each characteristic can be positively influenced without negatively influencing another.

A big problem with the CPAM factor goals is that it is well known that subproblem solution evaluation is often very difficult, as subsequent decisions may make apparently poor solutions perfectly acceptable in the context of the whole design, and vice versa.

Besemer (2008) reports that the CPAM is very useful for conceptual design evaluation. Given the significant impact of that phase on the rest of the design process, and the probable impact of conceptual design on creativity (Shai 2007), it makes sense to use CPAM characteristics in early design phases. As the selection of novelty moves the design problem towards being non-routine, there will need to be some sort of planning activity, such as gradual problem decomposition, plan fragment execution, or “opportunistic organization” (Visser 2006).

Next we will consider each CPAM characteristic to try to establish what a computational design system might do to affect its value in the final product profile.

4.1. COMPUTATIONAL NOVELTY

How can a computational design system produce something ***original****?* For it to produce something “infrequently seen in the universe of products made by people with similar experience and training” it must not produce something frequently seen. Thus knowledge of existing products of that type is required. This is a comparison between what is (being) produced and knowledge of prior products, judged by considering the whole product. The degree of difference can vary from, for example, a car with 5 wheels, to a car with just a new paint color. Judging *that* there is a difference doesn’t seem hard, but estimating the *amount* on some scale needs to be studied. The nature of the difference must be isolated and characterized, then some sort of conceptual distance needs to be estimated.

To produce the difference at all, a system needs to make design decisions that lead to new concepts, new combinations of concepts, and new values. This could be done by expanding the available choices for one or more decisions, or selecting a choice that was not previously made in that context.

To be ***surprising*** the product must present unexpected information to the evaluator: i.e., to the computational design system itself. The key problem here is for the system to know what is 'usual', and therefore expected. The system needs strong expectation knowledge for products, with representations that show stereotypes and reasonable variations. This may require a way to indicate the degree of similarity/difference between two items or concepts.

Unusual combinations of concepts (e.g., balloons as table legs, or cups made of cloth) need to be produced, perhaps by analogical reasoning, in order to "surprise" the evaluator. I would expect this characteristic to occur more if the design was being generated with the goal of exploration (e.g., 'I wonder what happens if we design teapots to be more like sheep?').

Choices intended to be original and/or surprising cannot just be made randomly, as requirements need to be satisfied and the product must still be useful. The more "weird" the choices are, the more the system will need to prune the results of its actions. I would expect such as system to have some tolerance for the unusual (perhaps provided as an input, or as a 'degree of creativity' goal), but also to have an accumulated indicator of how unusual the current partial design actually is (novelty stackup), as this represents some sort of risk.

4.2. COMPUTATIONAL RESOLUTION

How can a computational design system produce something ***logical****?* If the product follows the acceptable rules for the discipline then it is logical. Having and using technological knowledge, explicitly or implicitly as rules, is what computational design systems do well.

To be ***useful*** the product must have clear practical applications. This issue is a little blurred by the question of what the starting situation is for the design process: i.e., where the design activity starts in the conceptual to parametric dimension (Brown 1996), and how strongly the activity is being driven by a purpose for the product (e.g., it may just be exploratory, intending to discover the limits of the design space). For now, we'll consider useful just to mean that it has some use.

As *most* products are designed to fill some stated purpose, often with a detailed functional description, being useful ought to be easy to satisfy. However, judging how well it satisfies the functional requirements during designing may be difficult without a clear mapping of function requirements to behavioral and structural requirements. Even then, the product will need to be simulated (a computational "mental" simulation) in order to assess how well it fits with requirements and intended use. Desired patterns of interaction between the product (and its features) and the environment (often a “user”) define the function of the product (Chandrasekaran & Josephson 2000; Brown & Blessing 2005).

To be ***valuable*** the product must fill “a financial, physical, social, or psychological need”. That is, if the intended function of the product is needed in many situations, by many people, leading to results that are of consequence, then it is more “valuable”. Besemer (2006) suggests that value estimates result from a cost-benefit analysis. Exactly how these can be assessed by a computational system is unclear, unless the system has considerable knowledge of the potential environments for the product. Products with added functionality (e.g., a knife with blade and bottle opener) may be rated more highly on a cost-benefit scale, but that one with a very large number of attachments (e.g., corkscrew, marlinspike, etc.) would have weight and size costs.

To be ***understandable*** the product presents itself in a self-disclosing way, which is considered to be ‘user-friendly’. A computational system might assess the product’s usability by looking at the patterns of interaction with a user and evaluating the physical and cognitive difficulty of each interaction (Persad et al. 2007), or it might use the known affordances of certain features of the product (Brown & Blessing 2005).

4.3. COMPUTATIONAL STYLE

To be ***organic*** the product must have a sense of wholeness or completeness about it, with all the parts working well together. Besemer (2006) refers to this characteristic in terms that are mostly about the product’s visible structure and materials. A computational system would need to assess the “flow” of the product’s geometry, the snugness of fit, and the “harmonious” relationship between the colors, surface finishes and materials used. This suggests the needs for some basic ability to make aesthetic judgments (Reich 1993).

To be ***well-crafted*** the product must have been worked and reworked with care to “develop it to its highest possible level for this point in time”. This refers mostly to the finish of the product: no rough edges, or scratches. As such it appears to be more about manufacturing than design. However, insomuch as these specifications appear in the design description this is something that a computational system can assess.

To be ***elegant*** the product must be refined and understated. Besemer (2006) refers to something being simple but powerful, with “little surface decoration”, and low visible complexity. Besemer (2008) suggest that elegant is close to “simple” in meaning, but existing work shows that simplicity, or complexity, is a complicated issue (Balazs & Brown 2001; Summers & Shah 2003). While the visible complexity could probably be estimated computationally, the ratio of “power” to simplicity should be included in the measure of this characteristic, where power is some measure of the amount or the complexity of the functionality being delivered.

Computational assessment of the Style factor focuses on the structure component of an SBF model, although the focus on visual appeal suggests that a ‘visual’ or ‘appearance’ dimension might be added to the SBF model, consisting of the structural properties that affect the user but don’t contribute to the intended functionality, or to the structure.

5. Other Creativity Measures

This paper’s focus on Besemer’s model is not meant to imply that other measures are of no use. Quite the opposite should be the case. In this short section we examine three other selected pieces of research that include measures that might be incorporated in a CDC to affect its output.

Christiaans (1992) carried out creativity assessment experiments using seven attributes: creativity; technical quality; attractiveness; interestingness; expressiveness; capacity to integrate form, function and construction; and degree of prototypicality. He concludes that “design experts are less able to distinguish between the different attributes” than those less expert. It’s interesting that Besemer (2008) reports seeing the same sort of product profile shapes regardless of expertise. Christiaans found a close relationship between creativity, interestingness, and attractiveness. Designs that were judged as more prototypical were seen as of higher technical quality, but received low creativity ratings. Surprisingly, subjects didn’t seem to have any trouble assessing creativity as a single attribute when assessing products.

Shah et al. (2003) propose four measures for the effectiveness of a designer’s idea generation (ideation): novelty, variety, quality and quantity. While their work implies that humans provide the various weights and scores needed to use their equations to calculate values for the measures, something similar might be incorporated into a CDC system. However, measuring conceptual ideas will be harder to automate. They note that the four measures should not be combined into a single effectiveness measure, so an appropriate mix would need to be determined for CDC systems.

Wannarumon et al. (2008) use “algorithmic aesthetics” to drive a genetic algorithm. They conclude that their system does creative design. It is certainly an excellent example of the use of regular assessment to affect design activity. They use aesthetic evaluation (fitness) models that are equations constructed from variables such as complexity, golden ratio, mirror symmetry and rotational symmetry. Clearly, something like these models might be used to assess Style.

6. Summary & Discussion

In this section we will first summarize some of the main points from the earlier discussion of the creativity literature. The paper concludes with a discussion of related ideas concerning the future of CDC systems.

6.1 SUMMARY

To a certain extent the main points raised by the creativity literature are not very surprising. There's general agreement that there are degrees of creativity, that its determination requires judgment, and that such judgment requires knowledge that has temporal scope. Mere novelty is not enough for creativity, and most agree that at least 'utility' should also be evaluated.

Most writers stress the need to 'break away' from well known methods and solutions, plus a willingness to be 'playful': perhaps by not following 'normal' methods, by dropping or relaxing constraints, or by making unusual associations. Such associations would foster analogical reasoning, another commonly discussed aspect. As these, and other, modes of reasoning will tend to generate more possibilities than normal, the role of evaluation becomes even more important. Most writers seem to agree that such evaluation is highly knowledge-based.

In addition to associating concepts, and even blending them, the ability to avoid being tied to existing conceptual organization is important for creativity.

Most of these issues relate to CPAM's 'original', 'surprising' and 'useful' characteristics; i.e., they are mostly about novelty. The other characteristics raise more subtle and difficult issues. This suggests that at the very least, a detailed study of how these characteristics can be affected by actions within a CDC system is a way of pulling the study of CDC systems away from the focus on novelty.

6.2 DISCUSSION

CDC systems are going to be complicated, especially if we wish to *study creativity*. Although 'what' a CDC system can do is much more important than what technology to use, the need for multiple types of knowledge-intensive reasoning at different levels primarily suggests using agent technology with a variety of symbolic knowledge representations.

Agents have been used for design in Multi-Agent Design Systems (MADS) (Lander 1997), but the allocation of functionality to agents and the MADS architecture imposed is going to be critical (Dunskus et al. 1995).

For example, Minsky's proposal of Critics and Selectors is similar to the Sponsor-Selector mechanism explored by Punch et al. (1995), suggesting the use of specialized agents that have restricted functionality and knowledge, and using specialized knowledge structures (Berker & Brown 1996). In addition various kinds of capable evaluators will be essential.

We aren't advocating collections of agents as models of human design teams, as that adds a layer of coordination and communication to design (Visser 2006, p.199) that will 'muddy the water' for the study of creativity. However, agents can still have different goals and knowledge, leading to the kind of conflict that failing conceptual combination might produce, or the kind of memory activation that can be used to look for relevant knowledge or analogies.

Once CDC systems are in place a variety of experiments can be carried out to study the possible roles of different types of reasoning and knowledge. In addition, as there are many books and web sites dedicated to principles of creativity, as well as tools for defining problems, creating ideas, and selecting ideas (e.g., CreatingMinds 2007), it would be extremely interesting to use CDC models to explore the effect of these creativity enhancing strategies.

The agenda for CDC includes studying fairly fine grained computational creativity mechanisms, such as: detecting possible far analogies; combining mechanisms; adding mechanisms to *working* computational design systems; observing the changes in behaviour and results as systems are changed; and experimentation.

The field of computational design creativity must avoid the trap of assuming that design creativity is a single thing. Some past theories about designing have assumed that there is only one type of designing, for example, and have developed general models from specific examples. Designing by an individual varies depending on where in the conceptual-parametric, routine-nonroutine space the design activity starts. It varies depending on how purpose-directed the design activity is, or how much creativity is being rewarded (or tolerated). On top of these variations, the CPAM defines a multi-dimensional space for product creativity.

Another trap to avoid is just focusing on obtaining creative results, regardless of how they are obtained, as opposed to making hypotheses about and simulating the underlying processes. While systems such as those based on Genetic Programming are able to produce remarkable results (Spector 2008), it appears unlikely that they will tell us much about creativity. Even the commonly used list of 3-5 computational 'creative processes' can be seen as too coarse-grained, as many authors assume that just having these in a CDC is enough.

We'll need some form of explanation facility in the CDC systems, with explanations such as "I was influenced by …", "I was trying not to…", "it was analogous to…", or "I had a goal to…" Fine grained systems instrumented to report such things, because the reasoning used supports such explanations, should advance the CDC field.

Cross (2006, p.41) cites John Casti as concluding that despite their successes "we have learned almost nothing about human cognitive capabilities…from chess playing programs". It would be unfortunate if in the future people were saying the same thing about the area of computational creativity, and computational design creativity in particular.

Even if CDC systems are not yet feasible, systems could *aid* human designers, helping them to be more creative, instead of actually doing the creative designing. However, regardless of whether it provides evaluation, critiquing, or suggestions, the system would still benefit from a full or partial model of the design process in question, as well as a strong sense of the factors that influence product creativity assessment.

The use of CPAM factors and characteristics as goals to influence a CDC system and to evaluate its progress appears to be a promising idea, although it is clearly not without difficulties, and much more investigation needs to be done.

Acknowledgements

I would like to acknowledge Dr. Susan Besemer for her assistance and personal communication. Any misrepresentation of her opinions, or those of others, is the fault of the author.

References

Amabile, TM: 1983, The Social Psychology of Creativity, *Journal of Personality and Social Psychology*, 43: 997-1013.

Balazs, ME and Brown, DC: 2001, Design Simplification by Analogical Reasoning, in (eds.) Rizzi, Cugini & Wozny (eds.), *Knowledge Intensive Computer Aided Design*, Kluwer AP.

Berker, I and Brown, DC: 1996, Conflicts and Negotiations in Single Function Agent Based Design Systems, *Concurrent Engineering: Research and Applications*, 4(1): 17-33.

Besemer, SP: 2006, *Creating Products in the Age of Design*, New Forums Press, Inc.

Besemer, SP: 2008, Personal communication.

Besemer, SP and Treffinger, DJ: 1981, Analysis of Creative Products: Review and Synthesis, *Journal of Creative Behavior,* 15: 158-178.

Boden, MA: 1994, What is Creativity?, in MA Boden (ed.), *Dimensions of Creativity*, The MIT Press, pp. 75-117.

Brown, DC: 1996, Routineness Revisited, in M Waldron and K Waldron (eds.), *Mechanical Design: Theory and Methodology*, Springer-Verlag, pp. 195-208.

Brown, DC: 2007, *Some Reactions to Presentations at the 2007 Computational Creativity Workshop*, Technical Report, CS Dept., Worcester Polytechnic Institute, USA.

Brown, DC, Horner, R, Kim, M, Large, E, Liu, J, Meehan, E, Sloan, WN and Spillane, M: 1992, Experiences with Modelling Memory and Simple Learning in Routine Design Problem-Solving Systems, in M Green (ed.), *Knowledge Aided Design*, Academic Press, pp. 239-257.

Brown, DC and Blessing, L: 2005, The Relationship between Function and Affordance, *ASME Design Theory and Methodology Conference*, Long Beach, CA.

Chandrasekaran, B and Josephson, JR: 2000, Function in Device Representation, *Engineering with Computers*, 16, 162-177.

CreatingMinds: 2007, Wise ways and words in all matters creative, *http://creatingminds.org/index.htm*, accessed January 9, 2008.

Christiaans, HHCM: 1992, *Creativity in Design: The role of domain knowledge in designing*, Uitgeverij Lemma BV.

Cross, N: 2006, *Designerly Ways of Knowing*, Springer-Verlag.

Dasgupta, S: 1994, *Creativity in Invention and Design*, Cambridge University Press.

Dunskus, B, Grecu, DL, Brown, DC and Berker, I: 1995, Using Single Function Agents to Investigate Conflicts, *AI EDAM*, 9(4), 299-312.

Ekvall, G: 1995, Assessing the climate for creativity and change, *Proc. Conf. on Organizational Climate for Creativity and Change*, Center for Studies in Creativity, Buffalo, NY.

Erden, MS, Komoto, H, van Beek, TJ, D'Amelio, V, Echavarria, E, Tomiyama, T: 2008: A Review of Function Modeling: Approaches and Applications, in A Goel, R Davis & JS Gero (eds.), special issue on Multi-modal Design, *AIEDAM*, 22(2).

Eysenck, HJ: 1994, The Measurement of Creativity, in MA Boden (ed.), *Dimensions of Creativity*, The MIT Press, pp. 199-242.

Gero, JS: 1994, Introduction: Creativity and Design, in T Dartnall (ed.), *Artificial Intelligence and Creativity*, Kluwer Academic Publishers, pp. 259-267.

Gero, JS: 1996, Creativity, emergence and evolution in design: concepts and framework, *Knowledge-Based Systems*, 9(7): 435-448.

Gero, JS and Kannengiesser, U: 2007, Locating creativity in a framework of designing for innovation, in N Leon-Rovira (ed.), *Trends in Computer Aided Innovation*, Springer, pp. 57-66.

Goel, AK: 1997, Design, Analogy, and Creativity, *IEEE Expert*, 12(3): 62-70.

Grecu, DL and Brown, DC: 1999, Guiding Agent Learning in Design, in S Finger, T Tomiyama and M Mantyla (eds.), *Knowledge Intensive Computer Aided Design*, 237-250.

Horn, D and Salvendy, G: 2006, Consumer-Based Assessment of Product Creativity: A Review and Reappraisal, *Human Factors and Ergonomics in Manufacturing*, 16(2): 155.

Lander, SE: 1997, Issues in Multiagent Design Systems, *IEEE Expert*, 12(2): 18-26.

Lubert, T: 2005, How can computers be partners in the creative process, *International Journal of Human-Computer Studies*, 63: 365-369.

Minsky, M: 2006, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*, Simon & Schuster.

O'Quin, K and Besemer, SP: 1989, Development, Reliability, and Validity of the Revised Creative Product Semantic Scale, *Creativity research Journal,* 2: 267-278.

Persad, U, Langdon, P, Brown, DC and Clarkson, PJ: 2007, Cognitive Scales and Mental Models for Inclusive Design, *HCII2007: 12th Int. Conf. on Human Computer Interaction*.

Punch, WF, Goel, AK and Brown, DC: 1995, A Knowledge-Based Selection Mechanism for Strategic Control with Application in Design, Assembly, and Planning, *International Journal of Artificial Intelligence Tools*, 4(3): 323-348.

Reich, Y: 1993, A model of aesthetic judgment in design, *Artificial Intelligence in Engineering*, 8(2):141–153.

Shah, JJ, Vargas-Hernandez, N and Smith, SM: 2003, Metrics for measuring ideation effectiveness, *Design Studies*, 24**:** 111–134

Shai, O, Reich, Y and Rubin, D: 2007, Creative conceptual design: extending the scope by infused design, *Computer-Aided Design*, in press.

Shneiderman, B: 2007, Creativity Support Tools: Accelerating Discovery and Innovation, *Communications of the ACM*, 50(12): 20-32.

Smith, SM: 1995, Fixation, Incubation, and Insight in Memory and Creative Thinking, in SM Smith, TB Ward and RA Finke (eds.) *The Creative Cognition Approach*, MIT Press, pp. 135-156.

Smith, SM, Ward, TB and Finke, RA: 1995, Principles, Paradoxes, and Prospects for the Future of Creative Cognition, in SM Smith, TB Ward and RA Finke (eds.) *The Creative Cognition Approach*, MIT Press, pp. 327-335.

Spector, L (ed.): 2008, Genetic Programming for Human-Competitive Designs, Special Issue, *AI EDAM*, 22(3).

Summers, J and Shah, J: 2003, Developing Measures of Complexity for Engineering Design, *Proc. ASME DETC*, paper DTM-48633.

Talbot, RJ: 1997, Taking Style on Board, *Creativity and Innovation Management*, 6: 177.

Visser, W: 2006, *The Cognitive Artifacts of Designing*, Lawrence Erlbaum Associates.

Wannarumon, S, Bohez, ELJ and Annanon, K: 2008, Aesthetic evolutionary algorithm for fractal-based user-centered jewelry design, *AI EDAM*, 22(1): 19-39

Ward, TB, Smith, SM and Vaid, J: 1997, Conceptual Structures and Processes in Creative Thought, in TB Ward, SM Smith and J Vaid, *Creative Thought: An Investigation of Conceptual Structures and Processes*, American Psychological Association.

Ward, TB, Smith, SM, and Finke, RA: 1999, Creative Cognition, in RJ Sternberg (ed.), *Handbook of Creativity*, Cambridge University Press.

Weisberg, RW: 1999, Creativity and Knowledge: A Challenge to Theories, in RJ Sternberg (ed.), *Handbook of Creativity*, Cambridge University Press.

Wiggins, GA: 2003, Categorizing creative systems, *Proc. Workshop on Creative Systems: Approaches to Creativity in AI and Cognitive Science*, at the 8th Int. Joint Conf. on Artificial Intelligence, Acapulco, Mexico.