Title
As Useful as a Bicycle to a Fish: Exploration versus Constraint in Creativity

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN
1069-7977

Authors
Dix, Alan
McKnight, Lorna
Ormerod, THomas C.
et al.

Publication Date
2006

Peer reviewed

As Useful as a Bicycle to a Fish: Exploration versus Constraint in Creativity

Lorna McKnight (l.mcknight@lancaster.ac.uk), Thomas C. Ormerod (t.ormerod@lancaster.ac.uk),
Department of Psychology, Lancaster University, LA1 4YF, UK.

Corina Sas (c.sas@lancaster.ac.uk), Alan Dix (alan@hcibook.com)
Department of Computing, Lancaster University, LA1 4YW, UK.

Abstract

We report a study that examined the roles of exploration and constraint in design creativity. 48 student participants undertook mental and physical variants of a design synthesis task requiring the creation of objects from presented shape components. In contrast with previous research, creativity scores were not affected by a preinventive phase in which forms are assembled prior to the provision of a design goal. Instead, a preinventive phase led only to higher ratings of design originality and lower ratings of design practicality. No differences in creativity were found between mental and physical synthesis. The results suggest that a preinventive phase does not enhance the exploration of novel combinations but instead increases the degree of design constraint.

Introduction

Creative design has been identified as one of the key challenges for theories of human cognition (Simon, 1981), but progress in understanding its cognitive determinants has been slow. An agreed definition for creativity remains elusive, though it is generally accepted to involve the generation of ideas or artefacts judged to be both original and practical (e.g., Boden, 2004; Warr & O'Neill, 2005): a new design idea will not be judged as creative if it cannot be implemented or used for its intended purpose. Further difficulties arise in the measurement of creativity, since assessment of a creative act is inherently subjective. Measures may vary over time and between cultures. Moreover, controversy exists as to whether the focus in understanding creative ability should be on measuring individuals' traits (Marakas & Elam, 1997) or on processes and skills acquired through training and experience (Sternberg, Kaufman & Pretz, 2002).

Despite these difficulties, a consistent finding in laboratory studies of creativity (e.g., Smith, Ward & Schumacher, 1993; Wiley, 1998; Marsh, Ward & Landau, 1999) and realistic studies of design (e.g., Jansson & Smith, 1991) is that prior knowledge can inhibit idea generation. For example, Marsh et al describe as unconscious plagiarism participants' generation of stimuli that were clearly influenced by exemplars available in the task environment, despite a requirement for novelty. An implication of this finding is that creativity will be enhanced by activities that encourage exploration of a new space of possibilities free from the influence of prior knowledge and expectations.

An influential cognitive model of the creative process that reflects the importance of unshackling exploration from prior knowledge is the Geneplore model of Finke, Ward and Smith (1992). This model proposes that successful creative idea generation consists of two critical phases: the generation of ideas, which invokes processes such as retrieval, association, synthesis, transformation, and analogical transfer; and exploration of ideas, which involves discovery and assignment of attributes, conceptual interpretation and functional inference. Finke et al refer to the products of these phases as preinventive structures, that is, ideas that can subsequently be refined into design outcomes through successive generation and evaluation cycles via the imposition of product constraints. The model is in some respects prescriptive, in that it presents an idealised model of how successful creativity might happen, with no inherent preconceptions or limitations about the types of creativity that it can accommodate.

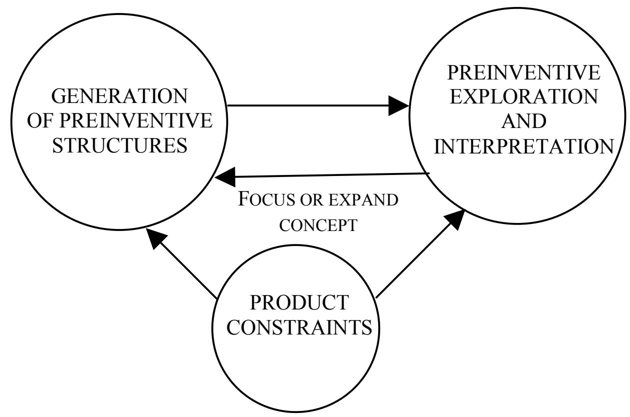

Figure 1: The Geneplore model of creativity (Finke, Ward & Smith, 1992).

Evidence for the Geneplore model comes from Finke's (1990) studies of mental synthesis. In this task, participants mentally generate and then report structures that address a design goal (e.g., a category that the resulting object should belong to or a function that it should serve) by combining a presented selection of generic three-dimensional shapes. The resulting products are then assessed by independent judges as to their originality and practicality. These measures are combined to produce a creativity score in which both originality and practicality must be above a high criterion. It is worth noting that the mental synthesis task, at least as used by Finke, does not test the cyclical nature of the model: in particular, typically in Finke's studies a single structure was developed during generation and evaluation, which was then interpreted as a design solution.

In a series of experiments, Finke demonstrated that creativity scores were enhanced if participants were instructed to synthesise a mental structure before knowing the design goal. This ‘preinventive phase’ was intended to force participants to follow phases of structure generation and subsequent interpretation. The preinventive structures should be free of interpretive bias when they were initially created, which should in turn increase the extent to which participants explore the space of possible interpretations. In these studies, generating preinventive structures before knowing the category or function under which the final product should fall was found to reduce the number of practical inventions, but produced the greatest number of creative inventions overall. However, these findings were based on cross-experiment comparisons rather than within-experiment tests.

Finke’s results have been very influential in design research (e.g., Kokotovich & Purcell, 2000), but to date no study has endeavoured to replicate his studies using the same methodology. The original studies are reported as book chapters, so only limited details about the methodology and materials are available. The products of the studies were rated by two judges who were both 'trained' (though no details are available concerning the training). No details were published about either the internal consistency of these judges, or the level of consensus between them. One aim of the current study was to provide as close a replication as possible of Finke’s (1990) second and sixth experiments, the critical comparison between presence and absence of a preinventive phase, given the available information on methodology.

According to the Geneplore model, a preinventive phase works because it allows participants to generate and explore shape combinations independent of prior knowledge about how they combine and operate in existing artefacts. However, an important aspect of Finke’s method was that, on each trial, participants in preinventive phase conditions combined shapes to synthesise a single form, to which the design goal was then applied. In contrast, participants informed of the design goal from the outset had the opportunity to synthesise and compare a number of potential combinations. If exploration lies at the heart of design creativity, one might expect the creativity of control participants to increase relative to preinventive phase participants because they had the opportunity to explore design alternatives. The fact that the opposite result arose in Finke’s studies suggests that exploration may help to construct novel pre-design combinations of components but not to generate design alternatives

An alternative explanation of Finke’s finding is that a preinventive phase acts as a constraint rather than as an opportunity to explore shape combinations. Preinventive form participants may be more creative because they have to fit an unpredictable design requirement to a predetermined form, which has the effect of increasing the originality of the resulting design. The decline in the number of practical inventions produced by preinventive phase participants is consistent with this view. If you are forced to re-conceptualise a fixed object in terms of an unexpected design goal, then it is likely to be highly novel at the cost of being unusable or ineffective: a bicycle makes a highly original fishing rod, but not a very practical one. A second aim of the current study was to provide evidence that might discriminate between explanations based on exploration versus constraint. If the former holds, then a preinventive phase should increase both originality and practicality. If the latter holds, then originality should increase and practicality should decline.

Although Finke’s studies involved three-dimensional shapes, the forms generated were imaginary rather than implemented. Typically, designers work with some external medium (e.g., sketching, building models). Thus, a third aim of this paper was to investigate whether the effects of a preinventive phase on physical synthesis are the same as on mental synthesis. Previous studies have examined physical synthesis, though these have not strictly replicated Finke's original experiments. For example, Pike (2002) found that a preinventive phase encouraged novel combinations of physical building blocks which in turn contributed to more creative products. However, in contrast to Finke’s studies, participants were informed that the study was about creativity, they were not blind to conditions, and had been lectured on the psychology of creative learning beforehand. Also, they were not required to use their preinventive structures in their final designs. Their study also used only self-reported measures of creativity and motivation.

Ishii and Miwa (2002) also explored physical manipulation in creativity, stating that "…the creative process should be characterized as both the external (physical) operation that tries to embody ideas into the physical world and the internal (mental) operation that considers ideas of products in the mind" (p178). Unlike Pike (2002), they report measures of originality and practicality, but physical synthesis was only used in the final building of products, not in preinventive exploration. They also used specialist materials (LEGO™ Mindstorms), which have inherent properties that might constrain creativity of the solutions.

Theoretically, one might expect differences in creative performance when participants use physical media. Various accounts of ‘distributed cognition’ (e.g., Zhang & Norman, 1994; Hutchins, 1995) suggest that features of a task can be represented and/or processed either internally in a person’s mind, or externally in the physical or visual task environment, taking the view that human knowledge is embodied not only in individuals but in objects and tools as well. On the one hand, physical objects might allow new combinations to be discovered by chance combinations that are hard to conceive mentally. On the other hand, physical media bring with them constraints like gravitational force, composite structure and textural cues that might limit creative performance. Noguchi (1999) found that the type of materials given to participants could affect their reported creativity, depending on the constraints inherent in the materials, although in this case the materials given had to be used in the finished design, rather than being representative of any material. In further support of this theory, research on the impact of sketching on creativity has shown that externalisation of a task through drawing can aid novel combinations of objects and interpretations of designs. In particular, Verstijnen, van Leeuwen, Goldschmidt, Hamel and Hennessey (1998) found sketching to be necessary for restructuring designs: in its absence people were more likely to persist with their original ideas. However, they did not examine physical manipulation of objects used for the tasks.

Experiment

An experiment was conducted to provide a within-experiment replication of Finke's (1990) Experiments 2 and 6, and to examine differences between mental and physical synthesis. All conditions were kept as similar as possible to the available details of Finke's original method, with the only intentional variation being the physical synthesis trials. If the Geneplore model holds, then a preinventive phase should enhance creativity irrespective of how objects are generated. Alternatively, the presence of physical components might variously affect synthesis through enabling new ways of exploring combinations or through the imposition of additional constraints.

Method

Participants

An opportunity sample of 48 undergraduate and postgraduate students recruited by advertisement were paid £5 each to participate. Of these, 20 were male, 28 were female, and all were blind to conditions.

Design

The experiment had three factors: Phase (preinventive phase or control), Synthesis (mental or physical), and Rater (self, independent or informed raters). Phase was a Between factor, and Synthesis and Rater were Within factors.

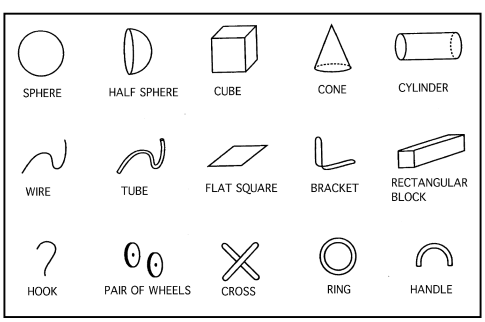

Figure 2: Object shapes (taken from Finke, 1992).

Materials

Shapes for each trial were selected quasi-randomly (using a computerised random number generator) from 15 basic 3D components (see Figure 2). Following Finke's method, the likelihood of selection was weighted so that generic shapes were selected most frequently. The shapes on the top row of Figure 2 (sphere, half-sphere, cube, cone, cylinder) had a 50% chance of being selected, the second row (wire, tube, flat square, bracket, rectangular block) a 33% chance, and the bottom row (hook, pair of wheels, cross, ring, handle) a 16% chance. For the physical manipulation conditions, shapes were made out of modelling dough. Flour-based dough allowed the shapes to be attached to each other, but did not harden. These shapes were made using moulds so that they were the same size and shape for all participants. Design categories were randomly selected from a list of 8 (shown in Table 1), with no category appearing more than once for a single participant.

Table 1: Design categories (taken from Finke, 1992).

	Category	Examples
1.	Furniture	chairs, tables, lamps
2.	Personal items	jewellery, glasses
3.	Transportation	cars, boats
4.	Scientific instruments	measuring devices
5.	Appliances	washing machines, toasters
6.	Tools and utensils	screwdrivers, spoons
7.	Weapons	guns, missiles
8.	Toys and games	baseball bats, dolls

Procedure

On each trial in preinventive phase conditions, participants were shown three shapes and had 1 minute to combine them to make "one single shape that does not represent anything in particular, but is merely an arrangement that you think is interesting and potentially useful in a very general sense". They then drew this form. Following this they were given a design category and spent a further 1 minute applying it to their design. They then wrote a brief description and labelled the resulting invention. Trials in the control conditions were identical except that participants were given the design category from the outset and had 2 minutes to conceive a design to meet the category before they drew, labelled and described their invention.

In Finke's study, four examples were shown to participants before trials began. These examples consisted of two 'uncreative' and two 'creative' inventions. Details of only one of each were published, so additional designs were created (exemplified in Figure 3) that were assumed to reflect the verbal descriptions given by Finke: a drafting pencil and wheelchair ('uncreative'), and a musical spheres game and an ice basher ('creative'). As in Finke's study, participants did not know they were taking part in a creativity experiment, nor were they told that they should try to be creative. Each participant received six trials, three with each medium. For each trial, a set of three shapes and a category were chosen. The same sets were replicated across four participants and the presentation order was counterbalanced across participants.

For trials requiring mental manipulation of the shapes, participants were presented with a printed sheet showing pictures of the shapes to be used (sampled from Figure 2). Where physical manipulation was required, on each trial three shapes made out of modelling dough were presented. Participants were told they might vary the size, position or orientation of any part in their design but should not bend or deform the parts, with the exception of the wire and the tube, which are bendable. The parts could be put inside one another; they might be hollow or solid, and in both the mental and physical conditions the shapes could be imagined to be made out of any material, including wood, metal, plastic, rubber or glass, or any combination of these materials. All three shapes were to be used, and, unless the same shape was specified more than once, only one should be used, with the exception of the wheels, which were in pairs. Participants were told that their designs need not be complete, and that they could imagine additional parts that would make it work such as motors, circuit boards and cogs.

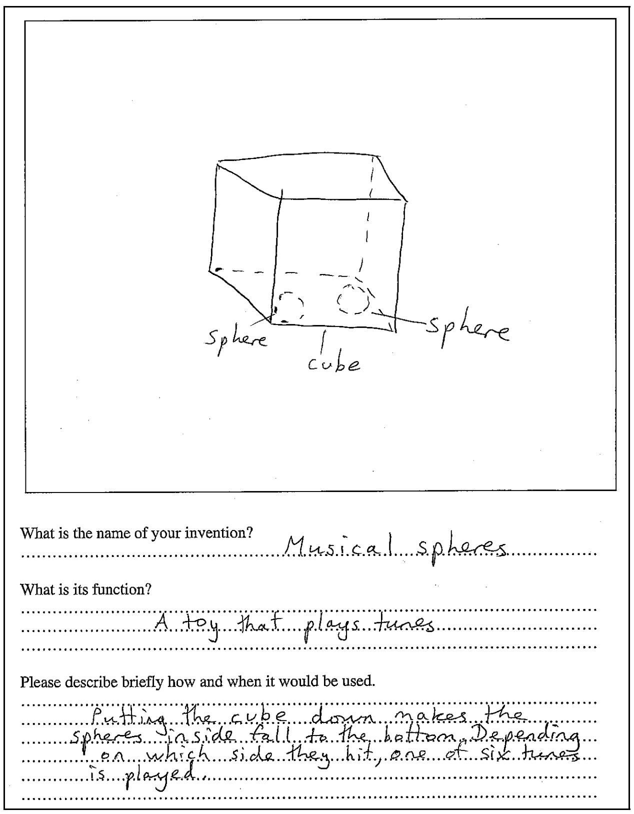

Figure 3: Design example: a musical spheres game (a 'creative' example).

Design Judging

After completing all the trials, participants were asked to rate each of their inventions on originality and practicality using Likert scales ranging from 1 (low) to 5 (high). They also rated how easy and how enjoyable they found each trial. Four independent judges also rated all designs for originality and practicality using the same scales. Judges were recruited in the same manner as the participants, and were paid £15 for their time. In addition, two informed judges also rated all designs by the same method, using the same criteria (an approximation to Finke's 'trained' judges). These judges were knowledgeable about the experiment, but for all judges the designs were presented randomly and they remained blind to conditions.

Each judge rated individually, with designs being shown in series on a computer screen. Judges were instructed to ignore the drawing skills of participants, and instead to base their judgements on the idea represented by each drawing and description. Judges were told to judge a design as 'unclassifiable' if no description was given or if they found the description uninterpretable. They were first shown the example designs (e.g., Figure 3) and then received five practice judgement trials. Independent judges were not told the experiment was a study of creativity or to rate the designs on how creative they were. The judging process was conducted in sessions lasting for at most an hour.

Results

Most participants completed the task without difficulty: participants failed to produce a design on only 21 out of 288 trials. Of these, 11 were failures to produce a design in the no-preinventive form condition, and 10 were failures to turn preinventive forms into inventions. These trials were excluded from the judging process. No participants in the preinventive phase condition were unable to create a preinventive form from the three shapes specified.

Finke reported that more 'highly creative' designs were found in the preinventive phase condition. Highly creative designs were ones that received a 'perfect score' of 10 (5 on both scales from both of Finke's judges). Designs with a score of 8 or higher were classed as creative. In the present experiment, however, no designs achieved an average score as high as 8 across all independent judges, the highest being 7.67. Although some designs did receive scores of 5 on both scales from some judges, there was never unanimous agreement. We chose to classify as 'highly creative' those designs that achieved a score above 7, which represents a score of at least 3 on each scale, with neither originality nor practicality being rated as low as 1. This only gives 22 out of 267 completed designs (8.24%), but there were marginally more of these designs occurring in preinventive phase conditions (13) than in control conditions (9).

Table 2 shows the mean ratings of originality and practicality for each rating group for mental and physical synthesis with and without a preinventive phase. Analysis of variance was conducted for originality and practicality measures between Phase, Synthesis and Rater factors. Analysis of variance was also conducted on the creativity scores, but no effects were significant. These scores represent the sum of originality and practicality, and an inspection of the means shows that summing these measures effectively cancels out any differences across conditions. Thus, we do not report creativity scores further.

An effect of Phase approaching significance was found in the originality ratings, F(1, 46) = 3.47, MSe= 1.25, p<.07. Designs in preinventive phase conditions (mean = 2.47) were rated as more original than designs in control conditions (mean = 2.23). A significant effect of Rater was found, F(2,46) = 32.95, MSe= .375, p<.01. Self-ratings (mean = 2.59) and independent ratings (mean = 2.53) of originality were higher than informed ratings (mean = 1.94). No other effects were significant with the originality measure.

Table 2: Mean ratings for each experimental condition

		Self ratings		Independent ratings		Informed ratings	
		Original	Practical	Original	Practical	Original	Practical
Preinventive phase	Mental	2.65	2.67	2.69	2.83	2.05	3.57
	Physical	2.88	2.83	2.57	2.84	2.01	3.74
Control	Mental	2.33	2.84	2.41	3.22	1.80	4.23
	Physical	2.49	2.78	2.43	3.09	1.90	3.75

A significant effect of Phase was found in the practicality ratings, F(1,46) = 4.38, MSe= 0.95, p=.04. Designs in preinventive phase conditions (mean = 3.08) were rated as less practical than designs in control conditions (mean = 3.32). A significant effect of Rater was again found, F(2, 46) = 70.91, MSe= .41, p<.01. Self-ratings (mean = 2.78) and independent ratings (mean = 2.99) of practicality were lower than informed ratings (mean = 3.82). There was also a significant interaction between Phase and Synthesis: mental synthesis (mean = 3.43) produced more practical designs than physical synthesis (mean = 3.21) in control conditions, while physical synthesis (mean = 3.14) produced more practical designs than mental synthesis (mean = 3.02) in preinventive phase conditions.

Self report ratings by experiment participants show that preinventive phase conditions (mean = 3.35) were rated significantly higher than control conditions (2.96) in terms of ease, F(1,46) = 5.45, MSe= .654, p = 0.02. However, there was no significant difference for enjoyment (preinventive phase conditions = 3.10, control conditions = 3.26). No effects of Synthesis were significant in either measure.

However, consistency between independent judges was not high, with a Cronbach's alpha score of 0.55 for originality ratings, 0.44 for practicality and 0.57 for the overall creativity scores. Looking at the raw data, it would seem that there was greater consensus on clearly unoriginal designs (including common objects such as lamps, chairs, cars and guns), but when the designs were more unusual, there was less consensus. Items that seemed highly original to some judges may simply have been uncommon items that those judges had not encountered before; other judges had (or believed that they had), and so gave them lower scores. There was also disagreement on the practicality scores: despite instructions, it seemed the judges had different internal measures of what constitutes a practical design. For example, some judges would rate a teleporter as being highly practical, as it would be extremely useful, whereas others would rate it as not practical at all, since it is infeasible using current technology.

The knowledgeable and independent judges rated designs in a very similar manner, although the knowledgeable judges gave slightly lower scores for originality on average than the independent judges across all conditions. Consistency between the knowledgeable judges was higher on originality and overall creativity, with respective alpha scores of 0.73 and 0.70, but only achieved a score of 0.37 on practicality.

Discussion

The results cast doubt on the previous findings of Finke (1990), and raise questions about the predictive validity of the Geneplore model. First, the ratings do not confirm the predictions of the Geneplore model. A preinventive form led to more original but less practical designs, as in Finke's experiments, but in contrast to his results, designs did not also differ in overall creativity scores according to whether they were produced after a preinventive phase or not. This outcome confirms the predictions of the alternative account based on constraint. In the mental synthesis paradigm, the requirement to impose a design goal on a single preinventive form limits the participant's ability to adapt a form to suit the design, so designs are more likely to be original. However, the imposition also makes designs less practical. In essence, the mental synthesis task does not address the cyclical nature of genuine design processes, in which adaptation and pursuit of alternative designs are fundamental to truly creative design expertise (Ball & Ormerod, 1995). The result is important because it challenges the widely held view that encouraging designers to explore the task environment prior to presentation of design goals enhances creativity: it may, but it is not tested by the mental synthesis paradigm.

The interaction between Phase and Synthesis is unexpected, but we believe it may be consistent with a constraint account of the preinventive phase. Much of the effect seems to come from the control conditions, in which designs produced mentally were judged as more practical than those produced physically. One possible explanation is that, when participants combine objects mentally to meet a design goal presented from the outset, their attempts are informed by prior knowledge of previous practical designs. Exploring design possibilities with reference to prior knowledge may be limited by physical building blocks that carry with them affordances or limits on feasible combinations. In other words, physical objects may limit the degrees of freedom available to participants to reproduce practical ideas, whereas mental objects allow exploration to enhance the practicality of the resulting artefact. In preinventive phase conditions, participants are forced to fit a design goal to a fixed shape combination, so prior knowledge of previous practical solutions to that goal cannot play a role in developing the shape combination. When the task of fitting a goal to a fixed shape is carried out mentally, there are fewer practical limits (e.g., gravity, friction) on how the fit is made than in physical conditions. Thus, our putative explanation for this interaction is based on the differential role across conditions played by prior knowledge and physical affordances in informing the practicality of resulting artefacts. This explanation is, of course, speculative and requires further empirical investigation.

A further concern about the validity of the mental synthesis task stems from the ratings procedure. Despite the increased rigour of the current experiment, the obtained ratings are far from convincing: inter-rater reliability was low even within rater groups, and rating groups differed still further. Practicality in particular appears to be a highly subjective measure. Moreover, our UK sample of raters appears generally much more conservative than the North American raters used by Finke. Future work with the mental synthesis paradigm requires more stable and objective measures than subjective ratings.

Finally, the study raises questions about the utility of definitions of creativity that hinge on a combination of originality and practicality. Previous work has sought definitions of creativity that assume additivity of these dimensions; in the current study however, they appear not to be additive but irrespective of conditions. Either the two measures should not be considered to be equally weighted in their importance to creativity, or the synthesis task fails to yield a significant number of creative inventions overall, which would seriously call the Geneplore model into question.

Acknowledgments

The study was conducted as part of the first author’s PhD research funded by a Doctoral Training award from the UK EPSRC.

References

Ball, L.J. & Ormerod, T.C. (1995). Structured and opportunistic processing in design: a critical discussion. *International Journal of Human Computer Studies*, 43, 131-151.

Boden, M.A. (2004). *The creative mind: myths and mechanisms.* London: Routledge.

Finke, R. (1990). *Creative imagery: discoveries and inventions in visualization.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Finke, R., Ward, T. B., & Smith, S. (1992). *Creative cognition.* Cambridge, MA: MIT Press.

Hutchins, E. (1995) *Cognition in the wild.* Boston: MIT Press.

Ishii, N., & Miwa, K. (2002). Interactive processes between mental and external operations in creative activity: a comparison of experts' and novices' performance. *Proceedings of the 4th Conference on Creativity & Cognition* (pp. 178-185). New York: ACM Press.

Jansson, D.G. & Smith, S.M. (1991). Design fixation. *Design Studies*, 12, 3-11.

Kokotovich, V. & Purcell A.T. (2000). Mental synthesis and creativity in design: an experimental examination. *Design Studies* 21 (5), 437-449.

Marakas, G.M. & Elam, J.J. (1997) Creativity enhancement in problem solving: through software or process? *Management Science*, 43, 1136 -1146.

Marsh, R.L., Ward, T.B. & Landau, J.D. (1999). The inadvertent use of prior knowledge in a generative cognitive task. *Memory & Cognition*, 27(1), 94-105.

Noguchi, H. (1999). How do material constraints affect design creativity? *Proceedings of the 3rd Conference on Creativity & Cognition* (pp. 82-87). New York: ACM Press.

Pike, C. (2002). Exploring the conceptual space of LEGO: teaching and learning the psychology of creativity. *Psychology Learning and Teaching*, 2(2), 87-94.

Simon, H. A. (1981). *The sciences of the artificial, 2nd edition*. Cambridge MA: MIT Press.

Smith, S.M., Ward, T.B. & Schumacher, J.S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, 21(6), 837-845.

Sternberg, R.J., Kaufman, J.C. & Pretz, J.E. (2002) *The creativity conundrum: A propulsion model of kinds of creative contributions*. New York : Psychology Press.

Verstijnen, I.M., van Leeuwen, C., Goldschmidt, G., Hamel, R. & Hennessey, J.M. (1998). Creative discovery in imagery and perception: combining is relatively easy, restructuring takes a sketch. *Acta Psychologica*, 99, 177-200.

Warr, A. & O'Neill, E. (2005). Understanding design as a social creative process. *Proceedings of the 5th Conference on Creativity & Cognition* (pp. 118-127) New York: ACM Press.

Wiley, J. (1998). Expertise as mental set: the effects of domain knowledge in creative problem solving. *Memory & Cognition,* 26 (4), 716-730.

Zhang, J. & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18*, 87-122.