Constructive Perception: A Metacognitive Skill for Coordinating Perception and Conception

Masaki Suwa (suwa@sccs.chukyo-u.ac.jp)
PRESTO, JST & School of Computer and Cognitive Sciences, Chukyo University,
101 Tokodachi, Kaizu, Toyota, 470-0348 Japan

Barbara Tversky (bt@psych.stanford.edu)
Department of Psychology, Stanford University,
Stanford, CA 94305-2130 USA

Abstract

One aspect of expertise, in chess, in diagrammatic reasoning, in design, and in other domains, is inferring conceptual ideas from perception. In design, this process is associated with creativity. We propose that underlying this skill is a process we call constructive perception, the deliberate adoption of perceptual strategies in the service of cognition. In the case of enabling new ideas in design, this seems to be a coordination of two processes: reorganizing perception and associating ideas. The present research presents evidence for the two components underlying constructive perception. Generating new interpretations of ambiguous sketches was correlated independently with a perceptual ability, reorganizing parts of figures, and with a conceptual ability, associative fluency.

Introduction

One common conception of creativity is facility at recombining old ideas into new ones. To this end, Finke and others have developed tasks that ask people to reconfigure parts for new uses (Finke, 1990). An analysis of this task reveals that it has two essential components: a perceptual component and a conceptual component. The perceptual component is reconfiguring visible parts. The conceptual component is molding the new configuration to some use, function, or goal. The perceptual component entails careful looking and mental transformation whereas the conceptual component entails connections to knowledge.

Viewed such, as a marriage of perceptual and conceptual processes, creativity comes to resemble expertise. For years artificial intelligence and cognitive science have addressed this issue. Expert systems that flourished in the 80’s in artificial intelligence were a bold trial to capture expertise in order to implement it in computers. But the inherent tacitness of expertise made it elusive to elicitation, a prerequisite for implementation. Direct pursuit of expertise was not a success.

A better understanding of expertise may come from investigating how the perceptual and conceptual processes of people become interwoven. Expertise embodies the kinds of features and relations in the external world an expert perceives and how the expert interprets them. This implies that experts are able to differentiate and perceive meaning in some features and relations in the external world that would be meaningless to novices or misinterpreted by novices. For example, the move history and options on a mid-game chessboard are not apparent to chess novices. Similarly, a statistical graph with two data lines that form an X may perplex a novice, but greeted with excitement by the expert, who sees in it an interaction. The ability to infer conceptual features from perceptual ones is often taken for granted by experts. Therefore, acquisition of expertise can be regarded as a process of becoming able to perceive what was not evident previously. Gibson and Gibson (1955) described a similar process, also in relation to expertise, in their case, wine-tasting: “Perceptual learning, then, consists of responding to variables of physical stimulation not previously responded to” (p. 34). Recent studies on the roles of external representations (e.g. Chandrasekaran, Glasgow & Narayanan, 1995) suggest that the interplay of perception and conception is a major driving-force of human problem solving and inference.

In the realm of design, expertise in perceiving plans and sketches is of particular significance. Sketches on paper are not simply instructions to engineers and contractors. Sketches also serve the designer as an external tool for checking the coherence and appearance of ideas as well as for generating new ideas and interpretations. Sketches are a revelation of a set of ideas as well as a stimulus for new ones (e.g., Goldschmidt, 1994; Schon, 1983). Arnheim (1969) called this “visual thinking”, providing many examples in art and architecture.

Reinterpreting sketches and generating new ideas work in a productive cycle. A detailed study of the cognitive processes of an expert architect revealed that when he made new perceptual discoveries in his own sketches, he was more likely to come up with new ideas. Similarly, new design ideas led him to see new features and relations in his sketches (Suwa, Gero and Purcell, 2000). That study highlights two significant components of expertise. One is perception of subtle features and relations in the external representation of a domain, and the other is the generation of ideas and
interpretations in the domain. Coordinating the two so that each amplifies the other is one important feature of both creativity and domain expertise, at least in design.

What does it take to be able to coordinate perceptual discoveries and conceptual ideas in a domain? Of course, exposure to many instances of perceptual discovery and conceptual generation through training in the domain is indispensable. Irrespective of training, however, some individuals are more effective than others. What underlies their skill? We propose that a mental skill, which we term, constructive perception, underlies coordinating perceptual discoveries and conceptual generation in a domain (Suwa and Tversky, 2002). Constructive perception entails awareness of the ways perception grounds interpretation as well as ways to reorganize perception in the search for new interpretations. We developed this concept to account for previous research observing students and designers in their attempts at generating as many interpretations as possible from ambiguous sketches (Suwa & Tversky, 2001) (see Figure 1 for the sketches). We found that professional designers produced more interpretations than design students, students who did not study design, or those whose profession was unrelated to design. This was a modest study, with 10 professional designers, 10 design students, 11 non-professionals and 22 non-design students. Why did professional designers perform better in this task, which is not directly pertinent to design? We proposed that the skill of constructive perception may be fundamental to expertise in design, and that it may be general enough to apply to other domains as well.

Here, we examine the abilities of designers in greater detail in order to gain a better understanding of the skill of constructive perception. What cognitive abilities are its significant components? In order to come up with a new interpretation of a sketch, it is helpful if not necessary to see the sketch differently, that is, to reorganize the parts or view it from a different perspective. But reorganizing is not sufficient; the new organization must be interpreted as well.

The evidence reviewed suggests that two abilities underlie constructive perception: a perceptual ability and an associative ability. Is the perceptual ability a general spatial ability or a more specific ability that engenders fluency in seeing new figures and relations? In the present study, we administer ability tests to expert designers, students studying design, people in unrelated professions and students studying cognitive science, and relate those ability measures to performance in a task requiring producing as many interpretations as possible for ambiguous sketches. The ability measures include a general test of spatial ability (mental rotation), a specific test of perceptual reorganization (embedded figures) and a test of (linguistic) associative fluency (remote associates).

**Constructive Perception for Interpreting Ambiguous Drawings**

Before describing the present study, we review the task of generating interpretations of ambiguous drawings, showing how the skill of constructive perception is required for good performance of this task. Reversing interpretations of ambiguous figures, such as the duck-rabbit figure, is difficult in imagery. Chambers and Reisberg (1992) argued seeing another figure requires changing perceptual reference frames -- the duck and the rabbit face opposite directions--, and that changing reference frames is more difficult in imagery. Rock (1973) has demonstrated that assigning a reference frame is integral to interpreting a figure. In general a drawing consists of elements arranged in space relative to each other and to a reference frame and perspective (Tversky, 2001). Interpreting a drawing means grouping certain elements and not others as well as assigning a reference frame and perspective. People unwittingly fixate to the particular groupings of elements and/or reference frames and perspectives underlying previous interpretations. This accounts for the difficulty, in general, of reorganizing perception.

Howard-Jones (1998) demonstrated fixation in an experiment on which ours was based. Participants were asked to generate as many interpretations as possible of a single ambiguous drawing. The typical pattern was a drastic reduction in the rate of generating interpretations after the first minute, the fixation effect. Fixation is a problem that plagues designers, indeed, all problem solvers. In order to avoid fixation, observers would need to coordinate two processes, perceptual reorganization of the elements of a drawing and conceptual generation of new interpretations. Perceptual reorganization can happen by regrouping elements, by changing reference frame, and by altering perspective.

We proposed that constructive perception is the skill for coordinating perceptual reorganization and conceptual generation. What is meant by “constructive perception” is self-awareness of the ways perception serves as the basis for interpretations of drawings. This awareness allows reorganization of perception in a deliberate and constructive manner, facilitating generation of new ideas and interpretations. Our previous study (Suwa, Tversky, Gero, and Purcell, 2001) using ambiguous drawings provided support for the hypothesis, showing that participants who were self-aware of how they perceptually grouped parts of a drawing and thereby regrouped parts intentionally generated more interpretations than those who did not. Perceptually regrouping parts of a drawing is one form of constructive perception. What the hypothesis implies is, in other words, that the performance of the task of generating many interpretations of an ambiguous drawing can be a measure of the skill of constructive perception.
Experiment: Comparing Generating Interpretations and Abilities

In this experiment, professional designers, design students, people in unrelated professions and students studying cognitive science were asked to generate as many interpretations as possible for four ambiguous figures. Their performance is compared to their performance on the three ability tests, perceptual reorganization, spatial ability, and associative fluency.

Methods

Participants. 48 people participated in the experiment. They consisted of 14 professional designers, recruited from the domain of industrial design or architecture; 12 design students, juniors or sophomores, 11 of which study industrial design at Tama Art University and 1 at Chiba University; 7 office workers in unrelated professions; and 15 students, all juniors, who study cognitive science at Chukyo University. The age range of professional designers was between 28 and 44, so that they have been practicing design for more than 5 years. The age range of office workers in unrelated professions was the same, between 28 and 44. They graduated from the department of law, social science, or humanity in their universities. Using participants from such diverse groups was expected to yield a wide range of scores for the task of ambiguous drawings.

Procedure. Participants first worked on the general spatial ability test, then the task of generating interpretations of ambiguous drawings, and then the perceptual reorganization test and the associative fluency test.

Ability Measures

We selected three tests as measures for the three abilities from the many measures that have been developed and tested (e.g. Carroll, 1993).

Mental Rotation. The most widely used measure of general spatial ability is mental rotation, because it correlates with performance in a wide range of tasks (e.g. Linn & Petersen, 1986). We chose to use the standard test of mental rotation, that developed by Vandenberg and Kuse (1978). In that test, participants judge whether a target stimulus is the same or a mirror image of a set of 4 test stimuli that are rotated relative to the target. There were two sets of 10 trials. For each trial, two of the four test stimuli are rotations of the target, hence, correct answers. The test is speeded; that is, performance is measured by the number of correct trials in a preset amount of time. Participants had 3 minutes for each set of 10 trials with a 1-minute rest between sets.

Embedded Figures. Perceptual redefinition or reorganization is one facet of “intellectual flexibility” (Guilford, 1959). The Embedded Figure Test (EFT) is widely used as a measure for testing perceptual redefinition. In EFT, participants are given many pairs of (a) simple figure(s) and a complex figure, and answer for each pair whether or not the simple one(s) is embedded in the complex one. Several versions of EFT have been proposed, varying in terms of whether the number of simple figures for each pair is one or more than one, and whether the number of simple figures embedded in the complex one for each pair, correct answers, is one (single-answer) or more than one (multiple-answer). Guilford (1959) argued that the multiple-answer condition of EFT requires the ability of redefinition more crucially. The reason is as follows. People, once they have found a simple figure to be embedded, would unwittingly see the complex one in a way that the simple one embedded appears evident. The multiple-answer condition requires checking if other simple figures are embedded. To do this, people would have to reorganize ways to see the complex figure, e.g. by regrouping its parts. For this reason we adopted the multiple-answer version of the EFT.

The material for the multiple-answer EFT was adapted from Gottschaldt’s test (1926), because the Gottschaldt figures have been most widely used (Guilford, 1959). There were 15 items, one per page. Each item had 5 simple figures and one complex figure. The simple figures were the same on each page. For 6 items, one simple figure is embedded in the complex one, for 6 items 2 simple figures are embedded, and for 3 items 3 simple figures are embedded (the Gottschaldt figures only allowed one correct answer; hence the adaptation). Participants were told that 1 or more of the simple figures could be found in the complex figure and that the orientation of the embedded figure is the same as in the simple figure but the size might be proportionally scaled up or down. Participants had 5 minutes to work on the task.

Associative Fluency. Associative fluency is defined as the ability to produce words in a restricted area of meaning. Kettner, Guilford and Christensen (1959) listed three tests having high loadings with this ability: Associations III, Associations IV, and Controlled Associations. The task in Associations III is to produce a word that is similar in meaning to two given words. The task in Associations IV is to produce a word that can be associated with two given words. The task in Controlled Associations is to write as many synonyms as possible for each given word. It is not evident from the previous literature which of the three has been most widely used. Our primary consideration in selecting a measure was appropriateness to Japanese because the participants in the experiment were Japanese. Japanese uses Kanji-characters each of which has inherent meaning(s). Taking advantage of this, Japanese language has a systematic way of creating a word,
called ‘jukugo’, by two Kanji-characters so that the word has the connotation of both meanings. If Japanese participants were asked to produce a word similar to a given word, it might be easy to produce a ‘jukugo’ that shares a Kanji-character with, and thus is similar in meaning to, the given word. Therefore, Associations III and Controlled Associations may be easy tasks in Japanese. For this reason, we selected Associations IV as a measure for testing associative fluency. We prepared 25 pairs of unrelated words in Japanese. For each, participants were to produce a third word that was associated to both, for example, producing “egg” as a third word associated with given words, “cake” and “frog”. Participants had 4 minutes to work on the task. Then, for each pair, they wrote the way each given word was associated to the produced word. These explanations were used for rating the plausibility of the produced words.

Task: Generating Interpretations

The procedure was adapted from Howard-Jones’s (1998). Participants generated interpretations for each of four ambiguous drawings, those shown in Figure 1, for four minutes, with the goal of generating as many interpretations as possible. For each four-minute session, participants had a stack of pages, each containing the drawing and a space for writing an interpretation. They studied the drawing, wrote an interpretation, put it aside, and repeated until the four minutes had passed. Participants were told that their interpretations did not have to include every part of each drawing. They were instructed not to rotate the sheet of paper.

Figure 1: Ambiguous drawings used

The drawings were systematically varied. The left two in Figure 1 used rigid lines; the right two used sketchy lines. Of each set, two contained primarily closed figures (the left of each set) and two contained relatively open structures (the right of each set). The order of drawings was the same for all participants: 1, 4, 3, 2.

Results

Scoring Ability Tests and the Generation Task

For Vandenberg test, we summed the correct responses over Parts 1 and 2 as the spatial ability score for each participant. For the multiple-answer EFT, we summed the correct responses over items without subtracting incorrect answers as the perceptual reorganization score for each participant. For the Associations IV test, we first rated the associations generated by participants. The principle adopted was whether or not both associations from the two given words to the produced word were understandable in terms of commonsense. Ratings were conducted blindly by the first author. The number of pairs for which both associations were rated as understandable was the score of associational fluency for each participant.

We reasoned that the more interpretations of ambiguous drawings a participant produced with a given time, the more skilled at constructive perception he or she is. Therefore, we adopted the total number of interpretations generated for the four drawings as a measure for the skill of constructive perception. There were no differences in number of interpretations for each of the four drawings.

Differences among the Groups

Did the four groups of participants differ in the three tests and the generation task? For the mental rotation test, one designer interpreted the task incorrectly, so the differences were examined over 47 participants. The one-way ANOVA indicated no significant differences among the four groups (F(3, 43) = 1.57). For EFT, another designer misinterpreted the task, so the differences were examined over 47 participants. The one-way ANOVA indicated no significant differences among the four groups (F(3, 43) = 0.962). For the associative fluency test, the designer who interpreted EFT incorrectly interpreted this task incorrectly as well, leaving 47 participants. The one-way ANOVA indicated significant differences among the four groups (F(3, 43) = 2.85, p<0.05). The designers had higher scores on associative fluency than the other groups, who were about the same. The post hoc Turkey’s HSD test indicated that for a significance level of 0.05, the critical HSD value was 2.69, so the differences between any pair was not significant.

For the task of generating interpretations, differences among the four groups were examined over 48 participants. The one-way ANOVA indicated significant differences among the four groups (F(3, 44) = 4.51, p<0.01). The post hoc Turkey’s HSD test indicated that the critical HSD value with the significance level of 0.05 was 19.1. The differences between the average number for designers (48.6) and that for cognitive science students (29.0), and between that for design students (49.2) and that for cognitive science students were larger than 19.1. This indicates that both designers and design students were more productive than cognitive science students.

The present results fail to replicate the previous finding (Suwa & Tversky, 2001) of higher performance of designers on generating interpretations than that of design students and office workers in unrelated professions. In the present experiment, both design students and office workers performed much better than
their counterparts in our previous experiment. The discrepancy might simply be variability due to small samples, or it might be the particular samples of designers and design students. However, the finding that a considerable number of students were able to perform in the generation task similarly to professional designers is worth noting. It again confirms the relevance of constructive perception to design. It also suggests that constructive perception is an ability on which individuals differ, and raises the possibility that students attracted to design are adept at it (or that students attracted to cognitive science are not).

Correlations of the Scores of Different Tests

More critical evidence for constructive perception comes from the correlations among the measures and with the generate ideas task. Do the scores of the three tests correlate with the total number of interpretations for the ambiguous drawings? Since two designers interpreted tests incorrectly, the correlations among the tests and the generation task were examined over 46 participants. Table 1 shows pair-wise correlations of the scores for the three tests and the generation task. Notably, there was no correlation between mental rotation score and numbers of interpretations; \( r = 0.065 \). The skill of constructive perception is not related to what is used as a general measure of spatial ability. On the other hand, there was a positive correlation between the score for perceptual reorganization, the multiple-answer EFT, and the number of interpretations; \( r = 0.253 \) (\( p < 0.05 \)). This finding adds to the evidence that spatial ability is not a unitary ability, but rather, one with several components (e. g., Linn & Peterson, 1986).

Table 1: Correlation coefficients of pair-wise examinations among the tests and the generation task

<table>
<thead>
<tr>
<th></th>
<th>Generation</th>
<th>EFT</th>
<th>Association</th>
<th>Rotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation</td>
<td>--</td>
<td>0.235*</td>
<td>0.358***</td>
<td>0.065</td>
</tr>
<tr>
<td>EFT</td>
<td>--</td>
<td>--</td>
<td>0.190</td>
<td>0.110</td>
</tr>
<tr>
<td>Association</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Note: the symbols ‘*’ and ‘***’ denote correlations at the 0.05 and 0.01 levels, respectively.

There was also a positive correlation between the score for associative fluency, the Associations IV test, and the number of interpretations; \( r = 0.358 \) (\( p < 0.01 \)). However, scores on EFT and on Associations IV were not related (\( r = 0.190 \)). This indicates that the abilities of perceptual reorganization and of associative fluency make independent contributions to constructive perception. In other words, it is not sufficient to be able to see new perceptual configurations or perspectives in external representations; the perceptual reorganizations must be linked to meaningful interpretations.

The independent contributions of perceptual reorganization and of associative fluency to constructive perception were corroborated by two additional analyses. As the first analysis, for each participant, we multiplied scores on EFT and Associations IV to form a composite score; the composite scores yielded an even stronger correlation with the numbers of interpretations than either score alone; \( r = 0.393 \) (\( p < 0.005 \)). As the second, we conducted multiple regression analysis. First, we set up the following multiple regression model;

\[
\text{Generation} = a + b \times \text{EFT} + c \times \text{Association} + d \times \text{Rotation} + \text{error}
\]

Fitting data for 46 participants to this model yielded the following set of coefficients; \( a = 21.19 \), \( b = 0.9109 \), \( c = 2.424 \), \( d = 0.0786 \). The statistical test showed that the model is not valid with the significance level of 0.05; \( F(3,42) = 2.77 \). The comparison of the three coefficients, \( b \), \( c \) and \( d \), reveals a less degree of contribution of the general spatial ability to the performance of the generation task than the other two abilities. Thus, we revised the regression model in the following way;

\[
\text{Generation} = a + b \times \text{EFT} + c \times \text{Association} + \text{error}
\]

For this model, the following set of coefficients were derived; \( a = 22.52 \), \( b = 0.9306 \), \( c = 2.425 \). Statistically, the model is valid with the significance level of 0.025; \( F(2,43) = 4.22 \). This confirms that perceptual reorganization and associational fluency are independent contributors to the generation task, but a general spatial ability is not. The analysis of validity of the coefficients \( b \) and \( c \) shows that \( b \) is valid, \( t(43) = 2.27 \) (\( p < 0.025 \)) and that \( c \) is marginal, \( t(43) = 1.35 \) (\( p < 0.1 \)). The contribution of associational fluency is stronger than that of perceptual reorganization.

Discussion

We began by observing that both a kind of creativity and domain expertise entail coordinating cognitive processes in two different modes; perception of subtle features and relations in the external environment, and conceptual generation of ideas and interpretations. We proposed that the skill of constructive perception may be fundamental to coordinating the two processes.

The present research examined constructive perception in a fairly general domain, interpreting ambiguous sketches. This has given support, in the domain of design, for the proposed idea by providing insight into the constituents of constructive perception, and suggesting how that skill may play a role in driving creative processes. Both associative fluency and perceptual reorganization contribute, and contribute independently, to constructive perception, supporting the idea that it has two components. Its perceptual component facilitates perceptually reorganizing what one sees in the external environment, enabling detection of subtle features and relations that novices might not discern. Its conceptual component enables fluency in generating new and related thoughts. To keep generating new ideas in a domain these two components must be tightly coordinated.

This analysis suggests a way that cognition is situated. Perceptual reorganization is grounded in external
stimuli. Expertise is attained by coming to differentiate and reorganize parts, wholes, perspectives, and reference frames. The perceptual expertise must then be coordinated with behavior, for example, playing a chess game or designing a building. For example, Clancey (1997) had claimed that coordinating perception, conception and physical performance is the fundamental way people and surrounding environments interact and develop together. The implication is that expertise is not acquired straightforwardly by teaching, but rather by a coordinated act of perceiving new features and relations unheeded previously, conceiving of new and associated thoughts unattended to previously and acting on the surrounding environment in a domain.

Constructive perception undoubtedly has some domain specificity. Reorganizing perception of chess configurations differs from reorganizer of checker boards or Go boards, and certainly from reorganizing sketches in design. Assigning interpretations to configurations similarly differs across domains.

Can constructive perception be nurtured? The previous finding (not replicated here) that professional designers were superior to design students and others in generating new interpretations is suggestive. The main thrust of the current research, that constructive perception consists of two components, perceptual reconfiguration and conceptual associations, suggests that promoting perceptual reorganization in coordination with interpretations might be effective. If such training succeeds, the next question is how domain independent the training can be and still be effective. A further question is whether analogous processes underlie creativity and expertise in other domains.

Acknowledgments
Preparation of this manuscript was aided by Office of Naval Research, Grants Number NO0014-PP-1-0649, N000140110717, and N000140210534 to Stanford University.

References

Kettner, N. W., Guilford, J. P. & Christensen, P. R. (1959). The relation of certain thinking factors to training criteria in the U.S. Coast Guard Academy, Educational and Psychological Measurement, 19, 381-394.