

# *Cognitive strategies of analogical reasoning in design: Differences between expert and novice designers*

Ozgu Ozkan and Fehmi Dogan, Faculty of Architecture, Izmir Institute of Technology, Urla, TR-35430 Izmir, Turkey

*This study investigates differences in analogical reasoning among first, second, and fourth year students and expert architects. Participants took part in an experiment consisting of four tasks: rating source examples, selecting a source domain, explaining their selection, and designing a bus stop. The results indicate significant differences among participants with respect to their soundness ratings. The results also show significant relation between level of expertise and participants' selection of source categories, the stated reasons for their selection, and the type of similarity they established between source and target. We conclude that experts preferred 'mental hops' while first year students preferred 'mental leaps.' Second and fourth year students preferred neither 'mental leaps' nor 'mental hops' but to literally copy the sources.*

© 2012 Elsevier Ltd. All rights reserved.

*Keywords: analogical reasoning, architectural design, creativity, design cognition, design education*

Analogical reasoning is described as a fundamental cognitive process underlying most other cognitive processes (Hofstadter, 2001), such as problem solving (Gick & Holyoak, 1980, 1983; Novick, 1988; Ross & Kilbane, 1997), scientific discovery (Dunbar & Blanchette, 2001; Gentner et al., 1997; Nersessian, 2008), learning (Brown, 1989; Vosniadou, 1989), and creativity (Johnson-Laird, 1989; Ward, 1998).

Gentner and Toupin (1986) claim that analogy is essential to both learning and discovery. Similarly, analogical reasoning plays a double role in design learning, supporting creativity and learning simultaneously. It is common among both architectural design students and practicing architects to browse through architectural publications in search of design examples that could be relevant to a design situation or for keeping up to date with recent projects. Furthermore, analogical thinking is a seminal learning strategy (Brown, 1989; Vosniadou, 1989) and enhances design learning. Design instructors often advise their students to enrich their visual vocabulary through studying masters' projects.

**Corresponding author:**  
Fehmi Dogan  
[fehmidogan@iyte.edu.tr](mailto:fehmidogan@iyte.edu.tr), [fehmidogan@gmail.com](mailto:fehmidogan@gmail.com)



[www.elsevier.com/locate/destud](http://www.elsevier.com/locate/destud)  
0142-694X \$ - see front matter *Design Studies* 34 (2013) 161–192  
<http://dx.doi.org/10.1016/j.destud.2012.11.006>  
© 2012 Elsevier Ltd. All rights reserved.

This study explores the role of analogy in design problem solving. It reports the results of an experimental study which investigated whether students from different levels of architectural education and expert architects differ from each other in the way they use analogy for a given design task. We specifically investigated the role of expertise in analogy, the impact of designers' goals on analogizing, and types of similarity established between sources and target domains. Compared to other similar research, this investigation included a larger and more diversified group of participants and used visual displays from four categories selected from a range of *near*, *near distant*, *medium distant*, and *distant* source examples. Furthermore, we identified the reasons for which subjects selected the specific source examples and source categories. The study makes a contribution to analogical reasoning in design by investigating the interaction among three factors: level of expertise, distance between source and target, and type of similarity established between source and target.

### *1 Analogical reasoning and design*

The use of analogy in design is common. Kalogerakis, Lüthje, and Herstatt (2010) found that analogies are widely used by professionals working at design and engineering companies. Ball, Ormerod, and Morley (2004) demonstrated that spontaneous use of analogy is common among both novice and expert designers.

Studies of analogy in design provide a wide range of results, yet are inconclusive in establishing a consistent pattern with regard to experts' and novices' preferences for near and distant domain analogies. Some researchers have discussed the differences between novices and experts primarily in terms of differences in their knowledge structure (Casakin, 2004, 2010; Casakin & Goldschmidt, 1999). Others have specified the differences more in terms of differences in goals (Ball & Christensen, 2009; Kalogerakis et al., 2010), in instructions to use analogy (Dahl & Moreau, 2002), or in the representational format of the source analogues provided to the subjects (Cardoso & Badke-Schaub, 2011; Zahner, Nickerson, Tversky, Corter, & Ma, 2010).

Kalogerakis et al. (2010) found that experts often prefer medium range analogies rather than near or distant domain analogies. Their results indicate that the preference between near and distant analogies depends on the goal of the designer. Similarly, Ball and Christensen (2009) claim that the distance between the selected source and the target relates to the purpose of the designer. In their study they found that expert designers use more between-domain analogies. In contrast, Christensen and Schunn (2007) found that experts use both within-domain and between-domain analogies, but primarily within-domain, when trying out improvements to existing products. Bonnardel and Marmèche (2004) found that experts evoke more between-domain sources and more sources in total; and when provided with between-domain source examples they are more likely to expand their search.

In a series of studies Casakin (2004, 2010; Casakin & Goldschmidt, 1999) investigated how experts and novices differ from each other in the way they use analogy. In the first study Casakin and Goldschmidt (1999) compared the design solutions and design ideas of beginning students, advanced students, and architects when given visual displays and instructed to use analogy. They found that architects and beginning students differed from each other significantly, whereas advanced students did not differ significantly from the other two groups. In the second study (Casakin, 2004), advanced students and architects were asked to use analogy and identify relevant sources from within-domain and between-domain examples. Novices identified significantly larger number of between-domain examples than within-domain examples; however, they retrieved between-domain examples as much as within-domain examples. Experts identified and retrieved significantly more between-domain examples. Experts and novices, however, were equally able to establish deep analogies. In the last study Casakin (2010) compared the use of analogy by architects, beginning students, and advanced students in well-defined and ill-defined problems. For well-defined problems the results indicated that when subjects were instructed to use analogy with visual displays available there was no difference in architects' and students' performance. When there were only visual displays and no instructions, however, architects performed better than beginning or advanced students. For ill-defined problems architects performed significantly better than novice students, both with explicit instructions and with no instructions, but there was no significant difference between architects and advanced students.

To encourage the use of analogies and avoid design or cognitive fixation many studies have manipulated either the given examples or the instructions to the subjects. One consistent finding is that when abstract representations are used as possible source analogues designers invoke more analogies and are better in analogizing (Linsey, Wood, & Markman, 2008). Bonnardel (2000) showed that when subjects are given names of objects rather than specific examples or no examples they spontaneously suggest more analogies. Others found that showing specific examples prior to a design or creative task often fixates subjects (Jansson & Smith, 1991; Marsh, Landau, & Hicks, 1996; Marsh, Ward, & Landau, 1999; Smith, Ward, & Schumacher, 1993). Cardoso and Badke-Schaub (2011) found that when subjects are shown examples that are too realistic they fixate more often. When abstract examples are shown the resulting designs are rated to be more original, which indicates less fixation (Zahner et al., 2010). Instructions to avoid irrelevant features of a shown source example seem to help avoid fixation as well (Chrysikou & Weisberg, 2005). Instructions to evoke as many analogies as possible also help designers (Dahl & Moreau, 2002) in proposing more original designs. Whether potential source examples are shown before or after the problem definition also impacts analogy in design. Tseng, Moss, Cagan, and Kotovsky (2008) found that if designers have an open-ended goal when

beginning the design process they are more likely to benefit positively from the given examples.

In this study we investigated the impacts of the level of expertise and the distance of the source domain to the target domain; we did not inquire about the impact of representational and instructional differences. Our first goal was to investigate whether participants with different levels of expertise are different in their preferences of source analogues, as suggested by previous studies (see [Ball et al., 2004](#); [Bonnardel & Marmèche, 2004](#); [Casakin, 2004, 2010](#); [Casakin & Goldschmidt, 1999](#); [Christensen & Schunn, 2007](#); [Kalogerakis et al., 2010](#)). Following the results of previous studies, our first hypothesis is that there will be differences among participants with different levels of expertise (Hypothesis 1). We used a  $4 \times 4$  design with four different groups of participants and four different groups of source domains to determine whether and how participant groups differ from each other with regard to the soundness ratings of given source examples. [Gentner, Rattermann, and Forbus \(1993\)](#) define soundness as the differentiation between poor and good analogies. We diversified the groups of participants to inquire how analogizing may change from novice to expert. First year students did not have any prior design education or experience, therefore, their performance during the experiment constituted a baseline. Similar to [Casakin \(2010\)](#), the student participants included not only first year students but also intermediary and advanced students. We wanted to investigate whether there is a change in the soundness rating of analogy for students throughout their education. Following Casakin's findings, we expected differences between beginning students and experts and we expected similarities between first year and second year students and between advanced students and experts.

[Gentner et al. \(1993\)](#) indicate that soundness ratings are related to relational and literal similarities rather than surface similarities. [Tourangeau and Sternberg \(1981\)](#) found that aptness in metaphors relate to distance between the source and target analogues, greater distance being positively related and less distance being negatively related to aptness. [Ward \(1998\)](#), however, claimed that in innovative analogies goals are the primary factors driving the retrieval process rather than distance. [Holyoak and Thagard \(Holyoak, 1985; Holyoak & Thagard, 1989, 1997\)](#) proposed that goals are one of the primary factors determining both retrieval and mapping in analogy.

We predicted two different possibilities for how participants would rate the soundness of source examples when asked to what extent the source analogues are useful. Following [Gentner et al. \(1993\)](#) and [Tourangeau and Sternberg \(1981\)](#), with distance as the main determinant of soundness ratings, we predicted that first year students would rate near-source examples higher because they lack sufficient domain knowledge to see the relational similarities with distant source domains. Experts, on the other hand, would rate distant source

domains higher because it would be easier for them to see relational similarities with distant domains. However, following Ward (1998) and Holyoak and Thagard (Holyoak, 1985; Holyoak & Thagard, 1989, 1997), if expertise level, together with particular goals of the participants, interactively impacts soundness ratings the results could indicate a more complex differentiation among participants.

If the difference between expert and novice groups of participants is related to difficulty in judging the distance between source domains and target domains, which might indicate differences in domain knowledge, one could help novices by providing more domain knowledge. If participants have a clear indication of the distance between source and target there may not be any differences in their soundness ratings to determine the usefulness of source analogues. We predicted that knowing the categories would help the participants assess directly the distance between the source and the target, therefore, assist novices significantly in rating a distant source domain. In other words, knowing the category membership might trigger top-down processing in analogical reasoning (see Ripoll, Brude, & Coulon, 2003), and, therefore, change participants' selection patterns. Our second hypothesis, therefore, is that when categories are disclosed to participants their differences will be less pronounced (Hypothesis 2).

## 2 *Mental leaps versus mental hops*

Different studies use different terminology to describe the distance between source and target, such as *within-domain* versus *between-domain* (Vosniadou & Ortony, 1989), *intradomain* versus *interdomain* (Holyoak, 1985), *intradesign* versus *interdesign* (Visser, 1996), *local* versus *distant* (Christensen & Schunn, 2007), and *near* versus *far* (Ward, 1998). All these terms basically refer to the same idea of distance between source and target. In the aforementioned studies this distance is a function of the degree of structural similarity (deep) and surface similarity (superficial) and establishes a continuum rather than a binary opposition (Dahl & Moreau, 2002). A structural similarity is about relational similarity, whereas a surface similarity is attributional (Vosniadou & Ortony, 1989). Different from these dichotomous categorizations, Dunbar (1997) proposed a tripartite categorization consisting of *within organism*, *between organisms*, and *nonbiological* when studying the use of analogies by molecular biologists. Following Dunbar's categorization Kalogerakis et al. (2010) proposed a tripartite scale (near, medium, and distant) to measure the distance between source domain and target domain in a design task. Their categories included *same product* category (near), *different product* category (medium), and *non-product* category (distant). Similarly, Ward (1998) proposed a tripartite classification, based on semantic similarity between source and target. According to Ward (1998), two analogues of a similarity relationship could be understood as belonging to either the 'same conceptual domain', 'related, though nonidentical domains', or 'wildly discrepant domains'.

In analogical reasoning genuine creativity is associated with establishing relationships to remote source domains that are difficult to bring to mind (Johnson-Laird, 1989; Ward, 1998), which Holyoak and Thagard (1997) name 'mental leaps'. Ordinary creativity, which Ward (1998) names 'mental hops', is associated with analogies to near-source domains and results in incremental innovation. Studying originality and its relation to the distance between source and target domains, Dahl and Moreau (2002) found that designers who invoked more remote source domains were more likely to produce original designs. Kalogerakis et al. (2010) found that distant domain analogies were associated with radically original designs whereas near-domain analogies were associated with incremental innovation.

When designers prefer near-domain analogies they are driven by factors other than originality. Near-domain analogies are cognitively less demanding and they offer a quick and efficient solution to a current problem. Ward (1994) identifies this preference as 'path-of-least resistance'. Near-domain analogies may lead to satisfying solutions when the goal is set accordingly, which may be the case for some design problems (see Christensen & Schunn, 2007; Kalogerakis et al., 2010). Kalogerakis et al. (2010) found that when economy of time is the utmost concern designers go for medium domain analogies, which balances time economy and originality. Similarly, Dunbar and Blanchette (2001) state that when scientists are simply trying to fix an ongoing experiment rather than set new hypotheses they go for within-domain analogies. The significance of goals in analogical reasoning is articulated in details by Holyoak and Thagard (Holyoak, 1985; Holyoak & Thagard, 1989, 1997) who claim that structural, semantic, and pragmatic considerations are of equal importance for analogy. Ward (1998) highlights the importance of goals also suggesting that for inventive analogies a match between source and target is possible only at the level of goal, since there is no target representation early in the creative process.

The second goal of this study is to inquire how and whether specific goals set by designers impact their source domain selections and whether there are differences among groups in their specified selection rationales. Here we expected to find out the underlying reasons for which particular source domains were selected. We predicted that there would be differences in the stated criteria with respect to levels of expertise (Hypothesis 3) and that the results would tell us more about the participants' reasons for selections and soundness ratings. We expected that those participants who selected near-domain sources would have done so because their goal was either cognitive economy (Ward, 1998) or time economy (Kalogerakis et al., 2010), the goal of those who selected distant domain sources would be originality (Johnson-Laird, 1989). Furthermore, we predicted that if goals were one of the primary factors in analogizing, then students would select analogies as distant as experts. We predicted that, regardless of differences in expertise levels, those participants

whose self-specified goal was originality would more likely select distant domains, whereas those whose goal was shaped by cognitive or/and time economy would more likely select near-domain categories.

### 3 *Difficulty in analogical reasoning*

The difficulty in seeing similarity relationships between remote domains is often associated with differences in expertise, either with reference to an underlying shift in the analogical reasoning process (Gentner & Toupin, 1986) or with reference to an ongoing change in mental representations (Vosniadou & Ortony, 1989). In either case, it is thought that experts in a domain have a structured and connected body of knowledge, whereas novices lack such a structured knowledge (Chi, Feltovich, & Glaser, 1981) and their similarity judgments are based on superficial features rather than structural features. Gentner and Toupin (1986) call the degree of relationship between surface and deep similarities 'transparency.' When surface similarities are directly indicative of deep similarities there is high transparency, and when there is no direct connection with deep similarities there is low transparency. Gentner and Toupin (1986) found that when there is high transparency between structural and surface similarity both young and older children have higher transfer accuracy. One possible explanation for the primacy of superficial features comes from Medin and Ortony (1989), who propose that there is often a non-arbitrary and causal relationship between surface and structural similarities. Holyoak and Koh (1987) relate difficulty in accessing remote domain analogues to the primacy of surface features. They claim that retrievals based on surface features alone will activate many competing associations among which it will be difficult to access the relevant remote analogues.

Most researchers agree that the distance between source and target domains is an important indication of the difficulty or ease of analogizing (Casakin, 2004; Christensen & Schunn, 2007; Herstatt & Kalogerakis, 2005; Holyoak & Koh, 1987; Johnson-Laird, 1989; Visser, 1996). Retrieval of near-domain sources is easier (Gentner et al., 1993) because they share a high degree of surface similarity with the target domain. Surface similarities are perceptual similarities which are cognitively less demanding than structural similarities, which are abstract. Keane (1987) found that subjects were more likely to retrieve literal analogues than remote ones. Novick (1988) found that both novices and experts had a default strategy to use surface similarities, but experts were more likely to avoid *surface similar yet structurally dissimilar* analogues. Holyoak and Koh (1987) found that both surface similarities and deep similarities have an impact on source retrieval; however, only deep similarities have a positive impact on the problem solution.

Beginning with the seminal work by Gick and Holyoak (1980) many researchers have focused on the conditions under which spontaneous retrieval or primed retrieval of relevant source analogues occur. Gick and Holyoak

(1980) found that when subjects were given a hint their problem-solving performance increased dramatically. In a second study Gick and Holyoak (1983) found that when two relevant sources and an associated diagrammatic representation were presented to subjects simultaneously their analogical performance increased significantly, compared to other conditions. Based on these findings Gick and Holyoak concluded that a successful retrieval and application of analogy requires a relevant schema induction by the subjects. Along the same lines, Clement (1994) found that when there are similarities between analogues at the schema level analogical performance is better than when the similarity is latent or manifest. These studies together suggest that successful analogies often require an alignment between analogues at a higher representational level in the form of abstract schemas and that experts will more likely establish deeper similarities between a source and target while novices will do the opposite.

Gentner and Markman (1997) propose a quadripartite classification of similarity between source and target domain, consisting of *anomaly*, *mere appearance* similarity, *literal* similarity, and *analogical* similarity, with increasing sophistication and difficulty from the first to the last. According to Gentner and Markman (1997), an *anomaly* indicates neither surface nor deep similarities between a source and a target; *mere appearance* similarity indicates more surface similarity than deep similarity; *literal* similarity includes those matches which share many deep and surface similarities; and *analogical* similarity includes those matches that share many deep similarities and few surface similarities. The difficulty in analogizing increases from attempting to establish mere appearance similarity to analogical similarity, because the latter requires the invocation of abstract schemas.

This brings us to our third goal in this study, which is to inquire into the relation between level of expertise and type of similarity established between source and target domain during analogizing. Our hypothesis is that there will be a relationship between expertise level and types of similarity established between source and target (Hypothesis 4). We predict that when analogizing experts will use more structural similarities but novices will establish more superficial similarities or anomalies because they lack higher-level abstract schemas. For second year and fourth year students we expect to find no clear preference for either type of similarity.

## 4 Study

### 4.1 Participants

Participants were 373 students from two schools of architecture ( $n = 172$  first year students;  $n = 118$  second year; and  $n = 88$  fourth year) and 30 expert architects with at least 10 years of experience in the field (*mean* age = 37.68; *SD* = 6.69). At the time of this study the first year students were either in the first or second



week of their architectural education, which suggests they had no experience in design. We recruited the student participants from two schools to minimize the impact of particular pedagogical approaches on analogizing.

In comparison to a similar study by Casakin (2004) we compared four different groups of subjects. We added one group, second year students, in addition to beginning (first year) and advanced (fourth year) students. Casakin (2004) compared the performance of beginning and advanced students to experts.

## 4.2 Tasks

At the outset of the experiment all participants were told that their general goal was to design a bus stop and that all four tasks in the experiment should be completed with that general goal in mind.

### 4.2.1 Task 1

Task 1 was devised to test Hypothesis 1, that there would be differences among participants with different levels of expertise with regard to the soundness ratings of source domains. For Task 1 the participants were asked to rate 40 source examples for the soundness of source example in the context of the bus stop design task and using a 1 to 5 rating scale (1 *Poor*, 2 *Fair*, 3 *Average*, 4 *Good*, and 5 *Excellent*). They were specifically told to evaluate the usefulness of each example as a source domain for designing a bus stop.

### 4.2.2 Task 2

Task 2 tests Hypothesis 2, that when categories are disclosed to participants the differences in their selections will be less pronounced. For Task 2 the participants were given the names of the four source example categories together with examples from each category. They were asked to select the one category which would best serve as an analogical source domain for designing a bus stop.

### 4.2.3 Task 3

Task 3 asked subjects to produce a written explanation of the reasons for their selection from Task 2. Students' explanations of their selections were coded according to an 8-category coding scheme by one of the authors. The eight categories were *function*, *form*, *symbolism*, *aesthetics*, *originality*, *design process*, *nature*, and *structure* (Table 1). A second, independent coder was asked to code 10% of the data for coding validation. The independent coder was instructed first about the coding scheme, then trained with a subset of the data before the final validation. There was 85% agreement between the two coders.

Task 3 was designed to investigate Hypothesis 3. This third hypothesis was that there would be differences in the stated criteria with respect to levels of expertise. We also expected to find that designers' individual goals would have an impact on their criteria for category selection.

**Table 1 Coding categories for Task 3**

<i>Categories</i>	<i>Explanation</i>	<i>Example</i>
Function	Any statement which relates to the way a bus stop will serve its users while they are waiting for a bus	Because climatic conditions have to be considered. The bus stop should be designed so that during winter one should not freeze from cold and during summer should not cook from heat. That is why bus stops need to be enclosed spaces.
Form	Massing and composition	I am of the opinion that the canopy and seats at the bus stop need to be considered in unity. That is why I am particularly attracted to holistic schemes with folding plates and shell designs.
Symbolism	The way the bus stop makes itself visible and discernible within the urban context through culturally significant shapes	I chose the mailbox image from the artifacts category. This image evokes a bus because of its shape. This is why I would use the mailbox for my bus stop design.
Aesthetics	About the visually pleasing sensation that the bus stop invokes	To make an association between a bus and the stop. It attracted me; it is more pleasing to the eye.
Originality	Novelty and difference	I chose Nature. Because the form of the mushrooms on the tree was appropriate for a bus stop. And it will be the first such design which has not been designed by anybody else.
Design process	How the particular selection will facilitate or hinder the design process	It feels like a product or a building designed by somebody else is too constraining in design. I feel that when I look at a design it will be difficult to distance myself from the impact and the domain of use of that design. But when I look at the examples from nature I feel freer. I think I could concentrate better on the idea in my head if I could focus on the particular thing from the nature.
Nature	Nature itself is sometimes given as a justification for the selection	I believe that the best designs in life could be found in nature, and things derived from nature have excellence and are both useful and aesthetic.
Structure	The way the bus stop is supported	One of the reasons is that its structure is simple and easily graspable.

#### 4.2.4 Task 4

To test Hypothesis 4 the participants were told to select a source example from one of the four source domains and design a bus stop using analogy to the source example selected. Hypothesis 4 suggested that there would be different types of similarity relationships between source and target in the designs of participants with different expertise levels.

All designs produced in Task 4 were categorized by three independent judges according to four types of similarity, which consisted of *analogical* similarity, *literal* similarity, *mere appearance* similarity, and *anomaly* (Gentner & Markman, 1997). Judges were given a detailed introduction to the type of similarities used in the study and were trained before the coding. Each participant's bus stop design and the source example selected by the same participant were presented simultaneously. We used the Delphi Method to determine the consensus among the judges. In the first round of coding there was

84% agreement between at least two of the judges. For the remaining source examples we conducted a second coding; this round yielded a complete consensus among the three judges.









































### 4.3 Materials

We used 40 source examples from four categories (Table 2). In comparison to similar studies which investigated the use of analogy according to the dichotomy of near versus distant only (see Casakin, 2004; Dahl & Moreau, 2002), we followed the study by Kalogerakis et al. (2010) and increased the number of source domain categories to four. The source analogues were selected from near, near distant, medium distant, and distant source examples, which correspond to bus stop, architecture, product design, and nature categories, respectively.

The examples shown to subjects were selected in a two-phase process. In the first phase we identified 320 source examples comprised of 80 examples from each category. Then we reduced the number of examples for each category to 20 by means of rating the soundness for each category. In the second phase three independent judges who were expert architects were asked to rate the remaining 80 examples from 1 to 5, given that the example would be used in the task of designing a bus stop. We used the Delphi Method to determine consensus among the judges. We considered all ratings with SD less than one as similar. In the first round of rating there was 86% agreement among the judges. For the remaining source examples we conducted a second rating which yielded a complete consensus among the judges.

In the experimental set-up we provided source examples to the participants following a reception paradigm as opposed to a production paradigm. In studies of analogical reasoning following a reception paradigm source analogues are given to participants and in studies following a production paradigm participants are told to produce the source analogues (Blanchette & Dunbar, 2000;

**Table 2** The source examples grouped according to the four source categories

	1	2	3	4	5	6	7	8	9	10
BUSSTOP										
ARCHITECTURE										
ARTIFACT										
NATURE										

Dunbar, 2001). Differing from similar studies (cf. Casakin, 2004), we asked subjects to rate the soundness of the shown examples. We did not ask subjects to memorize nor learn the source examples because we did not inquire about encoding differences between experts and novices.

#### *4.4 Procedure*

A pilot study was conducted to evaluate and adjust the study procedure and materials. At the outset of the study one of the authors explained to the participants what analogy is and presented examples of analogy from the field of architectural design. Following this introduction the participants were told what their general goal would be during the experiment and that there would be four subtasks during the experiment (detailed above). During Task 1 participants were given 40 source examples in random order in the form of visual displays (see Casakin & Goldschmidt, 1999). Student participants completed the task in an auditorium-type classroom in which the visual displays were projected on a white screen. Each student had a folding chair. Students were told to not discuss the experiment with their friends during the experiment. Experts, on the other hand, completed the task either by themselves or in groups of two or three in their offices while viewing the visual displays on a computer screen. There was one display for each source example and each was shown for 15 s. Task 1 was completed in 10 min. The sequence and the timing of the displays were controlled by the investigators. During Task 2 participants were given all source examples, grouped according to their categories, on four A4 size print-outs. They were asked to select the one category which best suited the task of designing a bus stop. For Tasks 3 and 4 subjects were given as many A4 size sheets as needed to complete the tasks. Tasks 2, 3, and 4 were self-paced and took approximately 30 min. In total, the experiment lasted 40 min.

### *5 Results and discussion*

#### *5.1 Task 1*

We used Cronbach's alpha to measure the reliability of the results from the 40-item rating task. The soundness ratings were found to be highly reliable (40 items;  $\alpha = 0.83$ ). The results from Task 1 are given in Table 3. To determine whether there were significant differences among participants' ratings with respect to their expertise level we conducted a multivariate analysis of variance (MANOVA) using scores for each visual display as the independent variable and expertise level as the dependent variable. The results indicate a significant difference in the ratings of the participants according to their expertise level (Wilks'  $\Lambda = 0.42$ ,  $F(120, 1.022E3) = 2.7$ ,  $p < 0.05$  alpha level), which confirms Hypothesis 1.

A pairwise comparison of participant groups indicates significant differences among the participants, with the exception of fourth year students and experts

**Table 3** Frequency distribution of participants' ratings for the source examples

	<i>Bus stop</i>				<i>Architecture</i>				<i>Artifacts</i>				<i>Nature</i>			
	<i>First year</i>	<i>Second year</i>	<i>Fourth year</i>	<i>Experts</i>	<i>First year</i>	<i>Second year</i>	<i>Fourth year</i>	<i>Experts</i>	<i>First year</i>	<i>Second year</i>	<i>Fourth year</i>	<i>Experts</i>	<i>First year</i>	<i>Second year</i>	<i>Fourth year</i>	<i>Experts</i>
1-Poor	0.09	0.08	0.08	0.05	0.08	0.05	0.08	0.08	0.08	0.15	0.20	0.16	0.05	0.12	0.20	0.18
2-Fair	0.18	0.19	0.21	0.21	0.19	0.21	0.21	0.23	0.21	0.25	0.30	0.28	0.16	0.25	0.26	0.25
3-Average	0.28	0.29	0.28	0.22	0.29	0.30	0.29	0.30	0.28	0.27	0.23	0.29	0.20	0.26	0.24	0.26
4-Good	0.26	0.29	0.26	0.27	0.29	0.29	0.29	0.28	0.26	0.25	0.19	0.19	0.28	0.25	0.22	0.24
5-Excellent	0.20	0.15	0.28	0.27	0.15	0.15	0.13	0.11	0.17	0.09	0.07	0.08	0.31	0.11	0.07	0.07

**Table 4** A pairwise comparison of participant groups' source example ratings

	<i>First year</i>	<i>Second year</i>	<i>Fourth year</i>	<i>Experts</i>
First year	1	0.000	0.000	0.000
Second year	0.000	1	0.000	0.000
Fourth year	0.000	0.000	1	0.091
Experts	0.000	0.000	0.091	1

(Table 4). The pairwise comparison suggests that fourth year students are more similar to experts in the way they rate the source examples.

As predicted, these results show that soundness ratings significantly differ according to participants' expertise levels. Contrary to our predictions, experts rated near-domain examples higher compared to distant domain examples, in this case product design examples and nature examples. First year students, as opposed to experts, rated distant domain examples higher compared to near-domain examples. Fourth year students are more similar to experts in their ratings, i.e., rating bus stop and architecture examples higher than product design and nature examples. Second year students differ significantly from all the other groups; this suggests that after one year of design education they have a different way of looking at source examples but that their view is not similar to fourth year students or experts. Furthermore, there is no statistically significant difference between fourth year students and experts, suggesting that three-year of design education brings a substantive change.

In comparison to results from other studies our results indicate that first year students are more likely to prefer distant domain source examples. This implies that their soundness ratings are not solely determined by their level of domain knowledge, as implied by [Gentner et al. \(1993\)](#) and [Tourangeau and Sternberg \(1981\)](#) and that there might be other factors impacting the ratings, such as specific goals that designers set for themselves, as suggested by [Ward \(1998\)](#) and by [Holyoak and Thagard \(Holyoak, 1985; Holyoak & Thagard, 1989, 1997\)](#).

In problem-solving literature some have suggested that experts are more likely to notice similarities to distant domain analogies. [Johnson-Laird \(1989\)](#) proposed that genuine creativity requires an analogical relationship between disparate and distant domains and that forming such a relationship would require expert knowledge. [Gentner et al. \(1993\)](#) found that soundness ratings depend on higher-order relational similarities while retrieval depends on superficial similarities. [Tourangeau and Sternberg \(1981\)](#) found that aptness ratings for metaphors are associated with more distant domains. Why then did first year students rate the distant domain source examples higher than the near-domain source examples? One reason could be that there are differences in domain knowledge among participants. The first reason suggests that first year students might not have been able to judge the distance between a source

domain and target domain because of deficiencies in their architectural domain knowledge, while the other participants had enough domain knowledge to judge this distance. We investigated this hypothesis through Task 2. Another reason could be that there are differences in the participants' goals for analogizing. This would suggest that the participants had specific goals which shaped their ratings of source examples. In Task 1 experts and fourth year students might have taken neither the necessary time nor the cognitive challenge to establish an analogical pairing between two distant domains and first year students might have preferred to seek originality, a hypothesis into which we specifically inquired through Task 3.

## 5.2 Task 2

In Task 2 the source categories were disclosed and participants were asked to select a source example category among the four available ones rather than rate randomly presented source examples. The results for Task 2 are shown in Figure 1. Chi-square test results indicate that there is a significant relation between expertise level and category selection ( $\chi^2(9, N = 396) = 30.828, p = 0.003, p < 0.05$ ).

A pairwise comparison among groups of participants indicates that there is a statistically significant difference between expert designers and first year students and between fourth year students and first year students at the alpha level (Table 5). Regarding to across group univariate comparisons, experts and first year students were most significantly different ( $\chi^2(3, N = 392) = 3.075, p = 0.008, p < 0.05$ ) and first and fourth year students were second most significantly different ( $\chi^2(3, N = 389) = 3.075, p = 0.045$ ). Second year students were not significantly different from any other groups.

The results from Task 2 are similar to results from Task 1. Fourth year students are not significantly different from experts, second year students are not significantly different from any other groups, and first year students

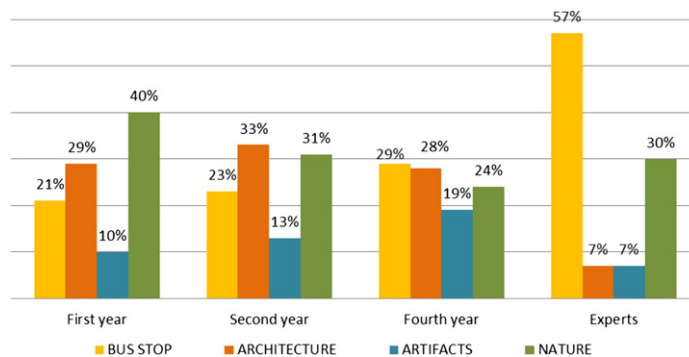


Figure 1 The percentages of category selections according to expertise level

**Table 5 Comparison of expertise levels for source category selection**

<i>Sig.</i>				
	<i>First year</i>	<i>Second year</i>	<i>Fourth year</i>	<i>Experts</i>
First year	1	0.221	0.045	0.008
Second year	0.221	1	0.425	0.066
Fourth year	0.045	0.425	1	0.208
Experts	0.008	0.066	0.208	1

significantly differ from all the other groups. When asked to select a category experts overwhelmingly selected near-domain source examples, whereas first year students chose mostly distant domain source examples. The differences in percentages of category selections for second year and fourth year students are less pronounced as compared to experts and first year students. The results indicate that second year students constitute an intermediary group between beginners and advanced students.

Results from Task 1 and 2 together indicate that students and experts rated source examples consistently within categories. Participants rated the source examples in Task 1 as if they knew about category membership. Our second hypothesis, therefore, was not confirmed. The participants did not change their preferences for source analogues even when they were given specific information about the distance between the source and the target.

We predicted that knowing the category membership might trigger top-down processing in analogical reasoning (see Ripoll et al., 2003) and, therefore, might change participants' selection patterns. In contrast to our prediction, even when categories were provided experts overwhelmingly selected near-source examples, while first year students chose distant source examples. Experts, therefore, did not prefer a distant domain source example even when they knew explicitly that it was further from the target. This suggests that experts may not always select a distant source domain in order to increase originality and creativity.

Based on these results we could propose that participants' ratings and preferences are not solely determined by their distance judgments but might also be shaped by specific goals. We investigated this hypothesis through Task 3. Perhaps experts take the path-of-least resistance (Ward, 1994) rather than undertaking a creative endeavor that is cognitively more demanding (cf. Dahl & Moreau, 2002). Furthermore, the results might suggest that without domain knowledge novices could base their selection on the expected originality of the end product without assessing what would be necessary to establish an analogical relationship between a pair formed of distant source and target domains.



### 5.3 Task 3

In Task 3 we asked participants to provide explanations for their selections of categories. Participants' explanations differed in the number of reasons they provided. Some gave only one reason or no reason while others gave up to four reasons. The coding for each participant yielded different numbers of categories; the percentages and ranking are given below (Table 6). We conducted a chi-square test to analyze the data and each time there was more than one coding for a participant each coding was included as a separate entry in the chi-square test analysis. A first chi-square test conducted according to expertise level shows that there is a significant relation between expertise level and participants' explanations ( $\chi^2(21, N = 579) = 66.281, p = 0.0001, p < 0.05$ ). A second chi-square test conducted with participants grouped according to their category selections yielded no results because of inappropriate expected frequencies. A post-hoc test comparing the standard residual values for each cell in the chi-square cross-tabulation to the critical value ( $-1.96$  and  $+1.96$ ) indicates five specific sources for differences among the groups of participants. First year students are significantly more likely to state aesthetic concerns (*std. residual* = 3.1) compared to all the other groups. They are also less likely to state formal considerations (*std. residual* =  $-2.0$ ) and structural ones (*std. residual* =  $-2.6$ ). Fourth year students are significantly less likely to state aesthetic concerns (*std. residual* =  $-2.4$ ). Finally, experts are significantly more likely to give design process-related reasons for their selections of source categories (*std. residual* = 2.9) when compared to the other groups.

For students the most commonly-stated reason for choosing a category related to functional considerations, while for experts it related to both functional considerations and to the design process. Among all the participant groups originality seemed to be most important for first year students. Often first year students expressed that the bus stop to be designed should be unique and different, i.e., original, and eye-catching or pleasing to the eye, aesthetically pleasing. When the percentages of aesthetics- and originality-related explanations are tallied for first year students they rank higher than functional

**Table 6 Percentages and rank order (in parenthesis) for coding categories according to level of expertise**

	<i>First year</i>	<i>Second year</i>	<i>Fourth year</i>	<i>Experts</i>
	% (Rank)	% (Rank)	% (Rank)	% (Rank)
Function	52.8 (1)	67.3 (1)	48.4 (1)	40.0 (1)
Design process	18.8 (3)	18.3 (4)	16.8 (3)	40.0 (1)
Form	13.2 (6)	26.0 (2)	23.2 (2)	16.7 (3)
Originality	18.8 (3)	8.7 (6)	10.5 (4)	10.0 (4)
Aesthetics	38.2 (2)	23.1 (3)	8.4 (5)	6.7 (5)
Nature	18.1 (5)	10.6 (5)	6.3 (6)	3.3 (7)
Symbolism	2.8 (7)	7.7 (7)	4.2 (7)	6.7 (5)
Structure	1.4 (8)	0.0 (8)	0.0 (8)	0.0 (8)

considerations. This suggests that first year students often have the goals of originality and being different. A first year student who chose the car example from the artifact group wrote that 'a bus stop in the shape of a car would be attractive because it has a different look.'

The results suggest that, while being concerned about functional issues, experts are equally concerned about strategizing the design process. Compared to the other groups they are more concerned about how their selections will shape the design process and about the effect of their selections on the expected quality of the end product. Based on these results we conclude that among participant groups there are differences in stated criteria for their source category selections, which confirms Hypothesis 3. We also conclude that first year students are more likely to name originality and difference as the reasons for their selection when compared to other groups, while experts are more likely to base their selection on design process-related reasons.

When all participants are grouped according to their selection of source example, however, we see similarities across groups (Table 7). Within all participants who chose *bus stop* examples as source examples and gave design process-related explanations for their selections (23 participants out of 84), 16 based their selection on the seemingly guaranteed success of the selection and the idea that a designer could improve a scheme based on a successful precedent to ensure either cognitive or time economy. Of all 84 participants who chose *bus stop* as source category only four stated that originality was the basis of the selection. Among those who chose examples from nature and gave process-related explanations (25 participants out of 128), 17 participants stated that the selection would potentially lead to a greater variety of ideas because of inspirational opportunities that the category provides. In contrast, only one participant within this group emphasized the guaranteed success of the selection when applied to a bus stop design task. Among all who chose examples from *nature*, 16 participants stated originality as the basis for the selection. Of 128 participants, 33 (25.8%) stated that examples from nature would

**Table 7 Percentages and rank order (in parenthesis) for coding categories according to source domains**

	<i>Bus stop</i>	<i>Architecture</i>	<i>Artifact</i>	<i>Nature</i>
	% (Rank)	% (Rank)	% (Rank)	% (Rank)
Function	60.7 (1)	70.7 (1)	54.3 (1)	44.5 (1)
Form	31.0 (2)	9.1 (5)	17.4 (4)	23.4 (3)
Design process	22.6 (3)	16.2 (3)	19.6 (2)	19.5 (5)
Symbolism	4.8 (4)	42.4 (2)	13.0 (5)	17.2 (6)
Aesthetics	4.8 (5)	6.1 (6)	8.7 (6)	1.6 (8)
Originality	27.4 (5)	10.1 (4)	19.6 (2)	21.1 (4)
Nature	2.4 (7)	0.0 (8)	2.2 (7)	32.0 (2)
Structure	2.4 (7)	0.0 (8)	0.0 (8)	0.0 (8)

lead to more creative designs either because of the originality of the end product or because of the fluency of the design ideas it would provide. When we look at the percentages of those who stated originality as one of the primary reasons for source category selection there is a considerable increase from *bus stop*, to *architecture*, to *artifact*, to *nature*. This suggests that when participants try consciously to maximize creativity they prefer distant domain analogies, whereas when they are after a tested solution to ensure the economic use of time or cognitive resources they select a near-source domain.

Some expert participants stated that using precedents would ensure the success of the end product. One expert, who chose the bus stop source category, wrote the following:

‘To solve the design problem, it seems that it is more appropriate to look at precedents that work. I think it is better to use tried and worked out solutions in problem solving. One has to be critical looking at precedents, yet material selections and form explorations would feed into the end design directly, since it is closer to the problem at hand. Perhaps one would have been more open-minded by looking at other category examples (such as artifacts, nature, etc...), but I personally think that we should focus on concrete architectural solutions since design is about providing a concrete and working product.’

Among students who chose bus stop source examples there are some who provided similar explanations. One first year student stated that choosing bus stop source examples would speed up the design process, provide better understanding, and lead to a better solution. A fourth year student wrote that because bus stop design follows determined standards and dimensions one should look at bus stop examples.

In comparison to participants who chose bus stop examples and gave design process-related explanations, those participants who chose examples from nature and gave design process-related explanations focused mainly on the creative potential of the examples. One first year student stated ‘predesigned objects restrain imagination while there are more inspiration and ideas in nature’. A second year student wrote that he chose the nature source domain because this ‘makes me feel freer and more comfortable about the bus stop ideas that are forming in my mind but not finalized yet’. An expert offered that choosing an example from nature will lead to a more unique solution.

Finally, very few participants (five in total) explicitly compared the advantages and disadvantages of different source domains and used elimination heuristics to decide on a source category. This group of participants compared either two or more than two source categories in terms of their idea-generation potential. For these participants the selection was more detailed and perhaps more conscious. A first year student wrote, ‘I didn’t select the first group [bus stop]

because I did not want to be influenced by them. It is difficult to find the solution in the nature group because this will require a wider perspective. I select the artifact group. It seems to be more creative.' An expert stated, 'I would make sure that the example I select comes from nature. To be inspired by other categories would be like creating a variation on them. Whereas to consider nature as the primary source offers more unique design opportunities.'

Based on the results of Task 3 we conclude that self-established goals are important in source category selection. [Ward \(1998\)](#) claims that in inventive analogies a match at the level of goals between a source and a target is the primary factor in source domain selection, since target domain is initially undetermined. Compared to analogical reasoning in other cognitive tasks, target in creative tasks becomes more and more determined during the course of analogizing. An initial match between source and target based on structural similarities, therefore, is impossible, which is why [Ward \(1998\)](#) suggests that in this type of analogies matching is goal-determined.

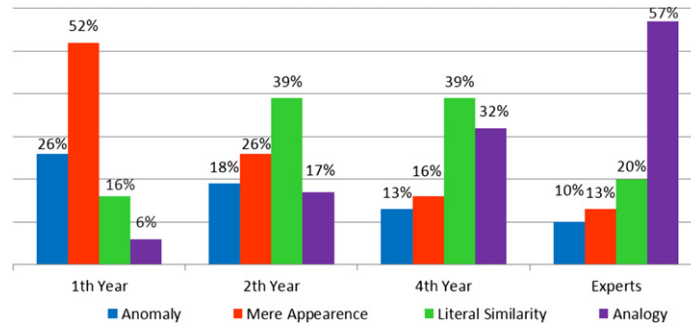
#### 5.4 Task 4

A match between source and target, however, does not seem to solely shape the solution process. [Chen \(1995\)](#) found that retrieval of a source analogue and also mapping depend on different types of similarities. [Holyoak and Koh \(1987\)](#) found that both surface similarities and deep similarities have an impact on source retrieval; however, only deep similarities have a positive impact on the problem solution. While retrieval depends on surface and structural similarities more, mapping depends on structural similarity and procedural similarity, i.e., the ability to see similarities in the steps to be taken when solving a problem.

Often, successful analogies require an alignment between analogues at a higher representational level in the form of abstract schemas during mapping. Experts will more likely establish deeper similarities between source and target with the help of abstract schemas while novices will do the opposite. When subjects lack such schemas or are not able to construct them they tend to follow verbatim the source analogues, and when they induce a schema they refer to the source analogue only occasionally ([VanLehn, 1998](#)). Similarly, [Ball et al. \(2004\)](#) found that expert designers are more schema-driven when using analogy, whereas novices are case-driven focusing on individual details of a case rather than its generalization across different cases. To investigate what participants do with their source selection during design generation we conducted Task 4.

In Task 4 we asked participants to design a bus stop using an analogy to a selected source example. Participants' designs were then categorized by independent judges in terms of anomaly, surface similarity, literal similarity, and analogy. The results ([Figure 2](#)) indicate that experts overwhelmingly

Figure 2 Type of similarity relationship between source and target with respect to level of expertise



established analogical relationships, i.e., structural similarities between source and target. First year students, in contrast, were much more likely to establish surface similarities as opposed to structural similarities. The differences in percentages were less pronounced for second year and fourth year students. Literal similarity relationships, consisting of both surface and structural similarities, are higher for both for second year and fourth year students. This indicates that with education and expertise, second and fourth year students access structural similarities between source and target yet cannot avoid using surface similarities in their designs.

Chi-square test results indicate that there is a significant relation between expertise level and type of similarity ( $\chi^2(9, N = 375) = 89.39; p = 0.000, p < 0.05$ ), which confirms Hypothesis 4. As repeatedly suggested in analogical reasoning studies, experts are more likely to assess the value of structural similarities and map them across domains; this is because of the structured nature of their domain knowledge (Chi et al., 1981). A second chi-square test questioning whether there could be a relationship between selected source domain and type of similarity yielded no significant relationship ( $\chi^2(9, N = 371) = 9.39; p = 0.402, p > 0.05$ ). This further supports the conclusion that the observed effect is more likely related to differences in expertise. We suggest that with increase in expertise level surface similarity and anomaly would decrease.

### 5.4.1 Design examples

In this section we provide seven examples of participants' designs to illustrate the kinds of similarity relationships participants established for the bus stop design task. These include one example for each type of similarity (anomaly, surface, literal, and analogical) together with one analogical example from a first year student and two analogical examples from near and medium distant source categories. We include an analogical example from a first year student in order to illustrate that first year students are sometimes, even though rarely, capable of establishing structural similarities with selected source examples. We include examples from near and medium distant domains to illustrate that establishing structural similarities with a source domain is not limited to only distant

domains. Our observation is that it is hard to detect features of source examples that are not discernible at first sight and that it is harder to infer new information based on what is visible. Both actions are common in experts but rare in novices.

*5.4.1.1 Anomaly: first year student.* In this example (Figure 3) a first year student is replicating the image of a bus stop with no concern for establishing any kind of similarity with the selected source example. The bus stop design example includes the basic features of any bus stop, such as horizontal and vertical panels to create an enclosure, with the addition of a sign which indicates the bus stop (here, in Turkish). These features are easily accessible from an individual's long-term memory without the presence of a source example.

*5.4.1.2 Surface similarity: first year student.* The second example (Figure 4) illustrates how a first year student tries to establish some relation with a source example but fails to go further than replicating surface features. Here the student draws a flat image of an eye without realizing the kinetic and spatial features of the human eye, such as a moveable lid and protected interior space. In the annotation next to the drawing the student acknowledges that s/he is trying to draw an eye rather than a bus stop.

*5.4.1.3 Literal similarity: first year student.* In this example a first year student is replicating the source example (Figure 5). The designer replicates the look of the source example and demonstrates understanding of how it could be used as a series of canopies to provide shelter for people waiting for a bus. Both the superficial and structural features of the source domain are used; however, these are not abstracted or adapted to the design problem.

*5.4.1.4 Analogical similarity from a distant domain: first year student.* Our data indicate that it is rare for novices to establish analogical similarities. There are, however, a few successful instances. In our fourth example



Figure 3 Example of anomaly. A first year student's source example selection from the architecture category (left) and the proposed design

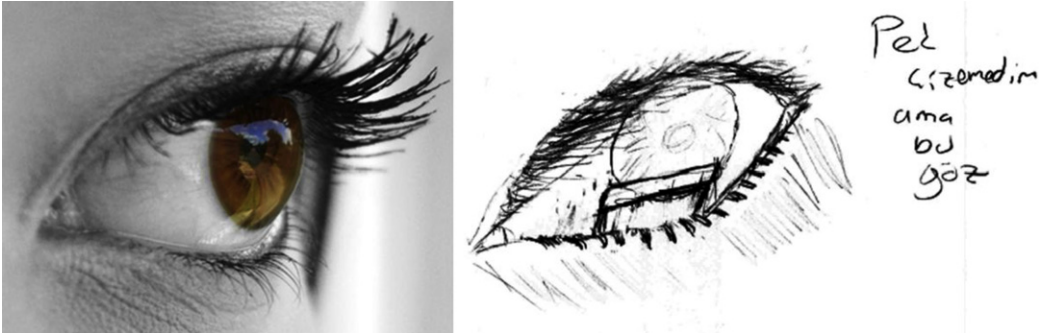


Figure 4 Example of surface similarity. A first year student's source example selection from the nature category (left) and the proposed design (right)

(Figure 6) we show a first year student's design for a multi-layered bus stop that is inspired by a human ribcage. There are some transfers of superficial features, such as the shape, yet the student demonstrates understanding of the general system of layering and the function of the central sternum as a unifying element in the overall structure.

*5.4.1.5 Analogical similarity from a distant domain: expert.* This expert participant's proposal is an example of analogical similarity (Figure 7). Here, the expert first analyzes the selected source example, a human eye, and concludes with an observation. The designer indicates that the eye consists of an inner shell and an outer shell which moves over the first. The designer shows how the two shells create a protected environment which changes



Figure 5 Example of literal similarity. A first year student's source example selection from the nature category (left) and the proposed design (right)

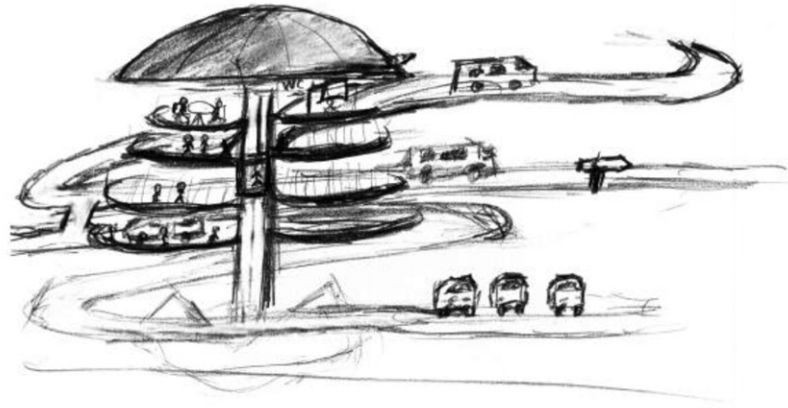


Figure 6 Example of analogical similarity. A first year student's source example selection from the nature category (left) and the proposed design (right)

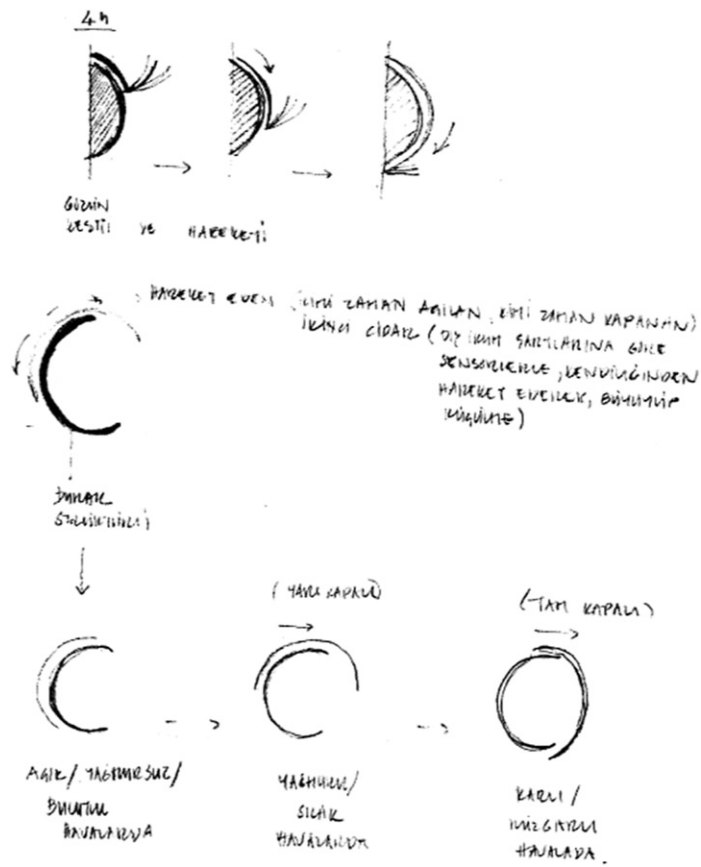


Figure 7 Example of analogical similarity. An expert architect's source example selection from the nature category (see Figure 4 left column) and the proposed design (right)



position according to weather conditions. In the first condition (fair weather) the outer shell is retracted; in the second condition (rainy or hot weather) the outer shell is halfway closed; and in the third condition (windy and cold weather) the shell is completely closed. In this example the designer manages to see what is beyond the immediately visible, abstract the salient features of the system, and adopt these features to a new design.

*5.4.1.6 Analogical similarity from a near domain: expert.* One might suggest that it would be harder for experts to avoid seeing the superficial features of a near domain. We include this example (Figure 8) to illustrate how an expert, while remaining loyal to the superficial features of a selected source, sees well beyond what is visible. The expert designer proposes that the bus stop example selected as a source could be interpreted as a modular system that might be enlarged from a single unit to multiple units according to needs. Here we see a clear attempt on the part of the designer to abstract the structural features of a source and adopt them for a generic design.

*5.4.1.7 Analogical similarity from a near-distant domain: expert.* In our final example (Figure 9) we show an expert using a source example (a kitchen cabinet set) from the domain of artifacts. Here the designer focuses on the mobility of each part of the cabinet set, derives the idea of a bus stop with movable parts, and proposes a ‘moving’ bus stop.

## 6 General discussion

When results from all four tasks are considered together we conclude that levels of expertise, distance between source and target, and the type of similarity established between source and target are three important factors that interact with each other in design analogy.



Figure 8 Example of analogical similarity. An expert architect's source example selection from the bus stop category (left) and the proposed design (right)



skills of abstraction, which would explain why they could only establish superficial similarities. Ball et al. (2004), for instance, found that in analogical reasoning expert designers are more schema-driven and novices are more case-driven, suggesting that experts would be more focused on higher-level abstract relationships whereas novices would tend to be more fixated on the specific details of a source example. As opposed to experts and first year students in our study, second and fourth year students were more likely to fixate. In contrast to first year students, fourth year students could see the full range of similarities between source and target and were, therefore, more likely to literally copy source examples. Compared with experts, fourth year students may not have yet developed an appreciation of incremental innovation which might help them avoid copying superficial similarities.

In contrast to experts, first year students select distant source domains more often because they primarily look for originality and difference. Novices could be following a simple heuristic, favoring originality judgments as opposed to efficiency and applicability. They lack, however, both domain and procedural knowledge with which they could establish structural similarities between distant source and target.

One of the obvious reasons why experts and novices differ during the design solution process is the novices' lack of domain and procedural knowledge, which they could use to construct abstract schemas and see deeper similarities between source and target. A second difference between the two groups relates to differences in the selection of source domains, which is more likely determined by goals. While experts set a goal of efficient planning of the design process, together with incremental innovation, the first year students set a goal of maximum originality (see Holyoak, 1985; Holyoak & Koh, 1987; Holyoak & Thagard, 1989, 1997; Ward, 1998).

In summary, first year students appear to seek 'mental leaps' without realizing what it would take to achieve them. Experts, in contrast, seem to prefer 'mental hops,' because they realize that radical innovation may not be efficient. The intermediary group does not seem to be seeking either 'mental leaps' or 'mental hops'. Instead they tend to literally copy the source examples.

## *7 Conclusion*

Our study results indicate that there is a significant difference among participants with respect to the way they rated source examples (Task 1) and that there is a significant relation between expertise level and source category selection (Task 2). Contrary to our predictions experts were more likely to rate near-domain examples and categories higher, whereas first year students were more likely to rate distant examples and categories higher. Knowledge of the category membership of each source did not affect these results. Experts gave more design process-related reasons for their selections of source

examples while the other participants gave more functional reasons for their selections (Task 3). When originality- and aesthetics-related reasons are tallied, the first year students' selections indicated that their dominant aspiration was to be original and different. Finally, experts were more likely to establish structural similarity, first year students established only surface similarity, and second and fourth year students established literal similarity (Task 4).

We conclude that experts were satisfied with 'mental hops'. In contrast to experts, first year students were trying to achieve 'mental leaps' but were only capable of establishing surface similarities. Second and fourth year students were not similar to either first year students or to experts in their source example selections and in the way they established similarities between the target and the source. They were inclined to literally copy source examples.

### *7.1 Limitations and future research*

We recognize that results from this study might have been different under different experimental conditions. Different types of instructions change analogical reasoning significantly (Casakin, 2010; Casakin & Goldschmidt, 1999; Chrysikou & Weisberg, 2005; Dahl & Moreau, 2002). If we had specifically instructed our participants to be innovative the results could have been different. As suggested by Holyoak and Thagard (Holyoak, 1985; Holyoak & Thagard, 1989, 1997) instructions matter because they set a particular goal that shapes the analogical process. Dahl and Moreau (2002), for instance, specifically asked subjects to use more than one source domain and the subjects produced more analogies. This is significant because the number of analogies invoked is positively associated with the originality of a design product. Another reason for our experts' preferences for source example categories might be the nature of the design task assigned; bus stop design may not be complex or challenging enough for expert designers to seek major innovations.

The format and number of visual displays provided to participants could have been manipulated also to examine in what ways realistic versus more abstract representations might impact analogical reasoning. The visual displays we provided were all photographs rather than abstract representations. Christensen and Schunn (2007) found that representations in the form of prototypes are more constraining than sketches. Some studies have found that abstract representations will less likely lead to fixation and constrain creativity (Zahner et al., 2010) and are more effective when used in problem solving (Gick & Holyoak, 1983). Similarly, we could have provided more than one visual display, using different formats.

We could also have changed the experimental set-up by asking participants to produce source analogues to be used in the later stages of the design and, therefore, follow a production paradigm in the study (Blanchette & Dunbar, 2000; Dunbar, 2001). Dunbar and Blanchette (2001) propose that when people

produce source analogues rather than receive them they are more likely to perform deeper analogies. Consequently, in-vivo studies of designers could lead to different results with regard to analogical reasoning. Studies of scientists working in a lab (Dunbar & Blanchette, 2001), of politicians and journalists (Blanchette & Dunbar, 2000, 2001), and of scientists from history (Gentner et al., 1997) indicate the prevalence and strategic use of deeper and within-domain analogies.

We identified self-established goals as one of the primary factors in participants' selection of source examples and categories. Another reason for the differences we observed among participants' design approaches could be the increase in conceptual knowledge of participants from first year to expert. Vosniadou (1989) proposed that differences between novices and experts and between children and adults in analogical reasoning do not relate to an increased ability to see structural similarities. Vosniadou suggests that structural similarities could become more accessible, or salient as people's conceptual knowledge of a domain becomes re-represented according to deeper principles. Once these similarities become more salient they may become the main trigger in analogical reasoning. We have not investigated, however, the full implications of Vosniadou's proposal. Further studies are needed to look into the role of changes in mental representations and their relation to analogizing in the field of design.

### *Acknowledgement*

We would like to thank Murat Gunaydin, Adile Avar, Serdar Kale, Ebru Cubukcu, Virginia T. Couch, and the anonymous reviewers who made valuable contributions during different phases of this study.

### *References*

- Ball, L. J., & Christensen, B. T. (2009). Analogical reasoning and mental simulation in design: two strategies linked to uncertainty resolution. *Design Studies*, 30(2), 169–186.
- Ball, L. J., Ormerod, T. C., & Morley, N. J. (2004). Spontaneous analogising in engineering design: a comparative analysis of experts and novices. *Design Studies*, 25(5), 495–508.
- Blanchette, I., & Dunbar, K. (2000). How analogies are generated: the roles of structural and superficial similarity. *Memory & Cognition*, 28(1), 108–124.
- Blanchette, I., & Dunbar, K. (2001). Analogy use in naturalistic settings: the influence of audience, emotion, and goals. *Memory & Cognition*, 29(5), 730–735.
- Bonnardel, N. (2000). Towards understanding and supporting creativity in design: analogies in a constrained cognitive environment. *Knowledge-Based Systems*, 13(7–8), 505–513.
- Bonnardel, N., & Marmèche, E. (2004). Evocation processes by novice and expert designers: towards stimulating analogical thinking. *Creativity and Innovation Management*, 13(3), 176–186.
- Brown, A. L. (1989). Analogical learning and transfer: what develops? In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369–412). Cambridge; New York: Cambridge University Press.

- Cardoso, C., & Badke-Schaub, P. (2011). The influence of different pictorial representations during idea generation. *The Journal of Creative Behavior*, 45(2), 130–146.
- Casakin, H. (2004). Visual analogy as a cognitive strategy in the design process: expert versus novice performance. *Journal of Design Research*, 4(2).
- Casakin, H. (2010). Visual analogy, visual displays, and the nature of design problems: the effect of expertise. *Environment and Planning B: Planning and Design*, 37(1), 170–188.
- Casakin, H., & Goldschmidt, G. (1999). Expertise and the use of visual analogy: implications for design education. *Design Studies*, 20(2), 153–175.
- Chen, Z. (1995). Analogical transfer: from schematic pictures to problem solving. *Memory & Cognition*, 23(2), 255–269.
- Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.
- Christensen, B. T., & Schunn, C. D. (2007). The relationship of analogical distance to analogical function and preinventive structure: the case of engineering design. *Memory & Cognition*, 35(1), 29–38.
- Chrysikou, E. G., & Weisberg, R. W. (2005). Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 1134.
- Clement, C. A. (1994). Effect of structural embedding on analogical transfer: manifest versus latent analogs. *The American Journal of Psychology* 1–38.
- Dahl, D. W., & Moreau, P. (2002). The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research* 47–60.
- Dunbar, K. (1997). How scientists think: on-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith, & S. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 461–493). Washington, DC, US: American Psychological Association.
- Dunbar, K. (2001). The analogical paradox: why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 313–334). Cambridge, Mass.: MIT Press.
- Dunbar, K., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: the case of analogy. *Trends in Cognitive Sciences*, 5(8), 334–339.
- Gentner, D., Brem, S., Ferguson, R., Wolff, P., Markman, A. B., & Forbus, K. (1997). Analogy and creativity in the works of Johannes Kepler. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 403–459).
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45–56.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: separating retrievability from inferential soundness. *Cognitive Psychology*, 25(4), 524–575.
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, 10(3), 277–300.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, 12(3), 306–355.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1–38.
- Goldschmidt, G. (2011). Avoiding design fixation: transformation and abstraction in mapping from source to target. *The Journal of Creative Behavior*, 45(2), 92–100.

- Herstatt, C., & Kalogerakis, K. (2005). How to use analogies for breakthrough innovations. *International Journal of Innovation & Technology Management*, 2(3), 331–347.
- Hofstadter, D. R. (2001). Epilogue: analogy as the core of cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 499–538). Cambridge, Mass.: MIT Press.
- Holyoak, K. J. (1985). The pragmatics of analogical transfer. In Gordon, H. B. (Ed.), *Psychology of learning and motivation, Vol. 19* (pp. 59–87). Academic Press.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15(4), 332–340.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13(3), 295–355.
- Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist*, 52(1), 35–44.
- Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, 12(1), 3–11.
- Johnson-Laird, P. N. (1989). Analogy and the exercise of creativity. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 313–331). Cambridge: Cambridge University Press.
- Kalogerakis, K., Lüthje, C., & Herstatt, C. (2010). Developing innovations based on analogies: experience from design and engineering consultants. *Journal of Product Innovation Management*, 27(3), 418–436.
- Keane, M. (1987). On retrieving analogues when solving problems. *The Quarterly Journal of Experimental Psychology Section A*, 39(1), 29–41.
- Linsey, J., Wood, K., & Markman, A. (2008). Modality and representation in analogy. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 22(02), 85–100.
- Marsh, R. L., Landau, J. D., & Hicks, J. L. (1996). How examples may (and may not) constrain creativity. *Memory & Cognition*, 24(5), 669–680.
- 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.
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York, NY: Cambridge University Press.
- Nersessian, N. J. (2008). *Creating scientific concepts*. Cambridge, MA, US: MIT Press.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 510–520.
- Ripoll, T., Brude, T., & Coulon, D. (2003). Does analogical transfer involve a term-to-term alignment? *Memory & Cognition*, 31(2), 221–230.
- Ross, B. H., & Kilbane, M. C. (1997). Effects of principle explanation and superficial similarity on analogical mapping in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(2), 427.
- Smith, S., Ward, T., & Schumacher, J. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, 21(6), 837–845.
- Tourangeau, R., & Sternberg, R. J. (1981). Aptness in metaphor. *Cognitive Psychology*, 13(1), 27–55.
- Tseng, I., Moss, J., Cagan, J., & Kotovsky, K. (2008). The role of timing and analogical similarity in the stimulation of idea generation in design. *Design Studies*, 29(3), 203–221.
- VanLehn, K. (1998). Analogy events: how examples are used during problem solving. *Cognitive Science*, 22(3), 347–388.

- Visser, W. (1996). Two functions of analogical reasoning in design: a cognitive-psychology approach. *Design Studies*, 17(4), 417–434.
- Vosniadou, S. (1989). Analogical reasoning as a mechanism in knowledge acquisition: a developmental perspective. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 413–437). Cambridge; New York: Cambridge University Press.
- Vosniadou, S., & Ortony, A. (1989). Similarity and analogical reasoning: a synthesis. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 1–19). Cambridge; New York: Cambridge University Press.
- Ward, T. B. (1994). Structured imagination: the role of category structure in exemplar generation. *Cognitive Psychology*, 27(1), 1–40.
- Ward, T. B. (1998). Analogical distance and purpose in creative thought: mental leaps versus mental hops. In K. J. Holyoak, D. Gentner, & B. N. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 221–230). Sofia, Bulgaria: New Bulgarian University Press.
- Zahner, D., Nickerson, J. V., Tversky, B., Corter, J. E., & Ma, J. (2010). A fix for fixation? Re-representing and abstracting as creative processes in the design of information systems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 24(2), 231–244.