

ON THE EFFECTIVE USE OF DESIGN-BY-ANALOGY: THE INFLUENCES OF ANALOGICAL DISTANCE AND COMMONNESS OF ANALOGOUS DESIGNS ON IDEATION PERFORMANCE

Joel Chan¹, Katherine Fu², Christian Schunn¹, Jonathan Cagan², Kristin Wood³, and Kenneth Kotovsky²

(1) University of Pittsburgh, USA, (2) Carnegie-Mellon University, USA, (3) University of Texas-Austin, USA

ABSTRACT

Design-by-analogy is a powerful method for innovation, particularly during conceptual ideation, but also carries the risk of negative design outcomes (e.g., design fixation, risk aversion), depending on key properties of analogies used. This paper examines how variations in analogical distance, commonness, and representation modality influence the effects of analogies on conceptual ideation. Participants in this study generated ideas for an engineering design problem with or without analogous example designs drawn from the U.S. Patent database. Examples were crossed by analogical distance (near-field vs. far-field), commonness (more vs. less-common), and modality (picture vs. text). For comparison, a control group generated ideas without examples. Effects were examined on a mixture of ideation process and product variables. The results show positive effects of far-field and less-common examples for novelty and quality of ideas; also, the combination of far-field, less-common examples increased novelty relative to control. These findings suggest guidelines for the effective use of design-by-analogy, particularly a focus on far-field, less-common examples during conceptual ideation.

Keywords: Design cognition, design methods, conceptual design, innovation, analogy

1 INTRODUCTION

Design-by-analogy has been shown to be a powerful tool for fostering innovation in engineering design [1-3], particularly during conceptual design, where design concepts are created either intuitively or through systematic processes. Analogy is a mapping of knowledge from one domain to another enabled by a supporting system of relations or representations between situations [3]. This process of comparison between situations fosters new inferences and promotes construing problems in new insightful ways; this in turn can be harnessed by designers to fuel innovative designs.

An illustrative and oft-cited example is George Mestral's invention of Velcro by analogy to burdock root seeds. The engineering design community has recognized the potential of design-by-analogy and developed a number of methods to utilize analogy within the design process; some examples include Synectics [5] – group design through analogy types; French's work on inspiration from nature [6]; Biomimetic concept generation [7] – a systematic tool to index biological phenomena that links to textbook information; the WordTree Method – a search hypernym and troponym search for analogies and analogous domains [8]; and analogous design using the Function and Flow Basis [9] – analogous and non-obvious product exploration using the functional and flow basis. These methods have been used with some success. However, fundamental questions surround the proper design and implementation of design-by-analogy methods.

It has been shown that exposure to analogous designs can result in lower levels of divergent ideation [10], and even inadvertent transfer of negative design elements [11]. These findings demonstrate that analogies, if not chosen or used well, can hinder rather than facilitate innovation; thus, understanding what makes for inspirational analogies is extremely important in order to effectively harness the power of design-by-analogy. To address this important gap in the literature, this paper examines the question of what characteristics of analogies, relative to the design problem, are most effective to improve ideation.

2 BACKGROUND

2.1. Analogical distance

Based on the psychological literature on analogy and problem-solving, analogical distance appears to be a key variable to consider. This variable can be conceptualized as ranging over a continuum from far-field (a different problem domain) to near-field (the same or very similar problem domain). Historical accounts of creative discoveries and inventions have often highlighted the potential of far-field analogies for creative insights, including George Mestral's invention of Velcro via analogy to burdock root seeds, and Niels Bohr's discovery of the structure of atoms via analogy to the solar system. Some cognitive studies in design have been similarly suggestive. For instance, the number of far-field analogies used by designers during ideation has been positively related to the originality of proposed solutions, as rated by a sample of potential customers [12]. Further, exposure to far-field design examples increase idea novelty relative to using no examples, and exposure to near-field examples decreases the variety of ideas generated relative to far-field examples [13].

On the other hand, far-field analogies can be difficult to retrieve from memory [14] or notice as relevant to one's target problem [15]. In addition, some investigators have disputed the privileged role of far-field analogies in prominent inventions and discoveries [16]. As such, it is possible that far-field analogies might result in some negative design outcomes. One way to tease apart possible ways in which far-field and near-field analogies might help or hinder designers is to use multiple measures of ideation processes, including novelty and variety of ideas, as well as average quality and variance in idea quality. An initial testable hypothesis is that providing far-field examples would allow one to generate more novel ideas relative to near-field or no examples.

2.2 Commonness

Another potential variable of interest is how common the analogous designs are found in designers' worlds. The commonness of a design increases the probability that a designer would have had prior exposure and/or experience with the design. This prior experience might influence the designer's ability to flexibly re-represent and use the design and combine it with other concepts in a novel fashion. This potential relationship between prior experience and flexibility of use is related to the psychological phenomenon of "functional fixedness," where individuals have difficulty seeing unusual alternative uses for common artifacts. For instance, in Duncker's [17] classic candle problem, the task is to fix a lighted candle on a wall in such a way that the candle wax will not drip onto a table below, and the given materials are a candle, a book of matches and a box of thumb-tacks. One correct solution involves emptying the box of tacks and using it as a platform for the candle; however, this solution eludes most solvers because it requires the unconventional use of the box as a platform. In fact, when the box is presented to solvers empty with the tacks beside it, solvers are much more likely to find the unusual solution [18]. Similarly, in Maier's [19] two string problem, where the task is to tie two strings together that are hanging from the ceiling just out of arm's reach from each other using various objects available (e.g., a chair, a pair of pliers, etc.), people often fail to recognize the solution of tying the pair of pliers to one string, swinging it like a pendulum, and catching it while standing on a chair between the strings.

Another potentially relevant finding in the psychological literature is that individuals who acquire experience with classes of information and procedures tend to represent them in relatively large, holistic "chunks" in memory, organized by deep functional and relational principles [20-21]. This ability to "chunk" supports expertise in routine, well-structured tasks [20], but might become a hindrance in tasks that require perceiving or representing information in novel ways (e.g., creative design), particularly if processing of component parts of the information chunks helps with re-representation [22].

These findings suggest that prior experience with analogous designs might hinder designers' ability to use those designs to fuel innovation. This leads to a hypothesis that less-common example designs, which designers are less likely to have been exposed to, might be more likely to support multiple interpretations, and thus facilitate broader search through the space of possible solutions, which might in addition increase the novelty of solutions.

2.3 Summary

Overall, the literature suggests that investigating variations in example analogical distance and commonness might shed some important light on the questions regarding what to analogize over. Prior work suggests hypotheses favoring far-field over near-field examples and less-common over more-common examples. Additionally, the theoretical and empirical literature suggest that there might be different effects of example analogical distance and commonness along different dimensions of the ideation process, thus motivating a fine-grained analytic approach to ensure that the effects of these variables can be clearly understood.

3 EXPERIMENTAL METHODS

3.1 Design

To investigate the effects of example analogical distance and commonness on conceptual design processes, we conducted a 2 (distance: far-field vs. near-field) x 2 (commonness: more common vs. less common) x 2 (modality: pictures vs. text) factorial experiment. The modality factor was included to control for specificity of effects to particular representation formats. In the experiment, participants, senior-level engineering students, were given a real-world design problem and were asked to generate solution concepts first briefly without examples, such that they understood the problem, and then with examples, to evaluate the effects of examples on problem solving. To establish whether examples of different types enabled or hindered problem solving, a control group of students executed a similar procedure but received no examples.

3.2 Participants

Participants were 153 students enrolled at two research universities in the U.S. 87% were undergraduate engineering students (95% Mechanical Engineering, 5% Electrical Engineering and others) and 13% Masters students in design-related disciplines (e.g., product development, business administration). Participants ranged from 20 to 38 years in age ($M = 22$, $SD = 1.89$), and were mostly male (~70%). Most participants had relevant mechanical engineering domain knowledge and design experience: 66% had at least 1-6 months of engineering internship experience, 82% had taken at least one course where a structured approach to design was taught, and 99% had experience with at least one prior design project in their engineering curriculum. Participants were recruited from classes and were given either extra credit or \$15 for their participation.

We randomly assigned participants to one of the nine possible conditions in each class by distributing folders of paper materials prior to students arriving in class. The obtained distribution of participants across the eight analogy conditions is shown in Table 1 (24 students were assigned to the control condition)—the sample populations, N s, are unequal not because of dropout, but rather from stochasticity in where students chose to sit down. With these sample populations, statistical power for detecting three-way interactions (not our theoretical goal) is modest, but power for detecting two-way interactions and main effects is good.

Table 1: Distribution of participants across analogy conditions

	Near-Field		Far-Field	
	More Common	Less Common	More Common	Less Common
Picture	13	17	15	16
Text	17	16	16	17

3.3 Design problem

The design problem was to design a low cost, easy to manufacture, and portable device to collect energy from human motion for use in developing and impoverished rural communities, e.g., rural India, many African countries. This design problem was selected to be meaningful and challenging to our participants. The problem was meaningful in the sense that real-world engineering firms are seeking solutions to this problem and the problem involves social value; thus, students would be appropriately engaged during the task. The problem was challenging in the sense that a dominant or accepted set of solutions to the problem has yet to be developed (so students would not simply retrieve past solutions), but it was not so complex as to be a hopeless task requiring a large design team and very detailed task analysis.

3.4 Selection of examples

Examples were patents retrieved from the U.S. Patent Database using keyword search on the U.S. Patent and Trademark Office website. Keywords were basic physical principles, such as *induction*, *heat transfer*, *potential energy*, as well as larger categorical terms like *mechanical energy*. The final set of eight patents was selected by two PhD-level faculty and one Ph.D candidate in mechanical engineering based on two sets of criteria: (1) balanced crossing of the analogical distance and commonness factors, such that there would be two patents in each of the four possible combinations, and (2) overall applicability to the design problem, over and above analogical distance and commonness. Each participant in the analogy conditions received two examples of a particular type, roughly balanced across conditions for applicability. The patents for each of the conditions are shown in Table 2.

With respect to the first set of criteria, in selecting for analogical *distance*, far-field patents were devices judged to *not* be directly for the purpose of generating electricity, while near-field patents were those judged to be directly for the purpose of generating electricity. In selecting for *commonness*, more-common patents were devices judged likely to be encountered by our target population in their standard engineering curriculum and/or everyday life, while less-common patents were those judged unlikely to be seen previously by the participants under typical circumstances.

Table 2: Patents for each condition

	Near-Field	Far-Field
More Common	-Waterwheel-driven generating assembly (6208037) -Recovery of geothermal energy (4030549)	-Escapement mechanism for pendulum clocks (4139981) -Induction loop vehicle detector (4568937)
Less Common	-Apparatus for producing electrical energy from ocean waves (4266143) -Freeway power generator (4247785)	-Accelerometer (4335611) -Earthquake isolation floor (4402483)

With respect to the modality factor, in the picture conditions, participants received a representative first figure from the patent, which typically provides a good overview of the device, while in the text conditions, participants received the patent abstract. To provide some foundational context, all text-and-picture-condition participants also received the patent title.

3.5 Experimental procedure

The experiments were conducted during class. Participants generated solution concepts in three phases, using a sequence of envelopes to carefully control timing of the task and exposure to examples across conditions. In particular, we wanted to ensure that design examples were received only after participants had made some substantial progress in ideation, since prior work has shown that examples and potential analogies are most helpful when received after ideation has already begun [23]. The overall time allowed for this task was sufficient to allow for broad exploration of the concept space, but not enough to develop particular ideas in depth, matching our focus on the ideation process.

Analogy and control groups executed the same overall sequence, but differed in the particular activities in the second phase of ideation. For the analogy groups, the sequence of phases was: (1) read design problem and generate solution concepts (10 minutes), (2) review two patents and write/draw solutions or ideas that come to mind when looking at the patents (10 minutes), (3) generate more solution concepts for the design problem (10 minutes), and (4) complete a brief design experience and demographics survey. For the control group, the sequence of phases was: (1) read design problem and generate solution concepts, (2) generate more concepts, (3) generate more concepts again, and (4) complete the design experience/demographics survey.

Participants were instructed to generate and record as many solution concepts to the design problem as they could, including novel and experimental ones, using words and/or sketches to describe their solution concepts.

4 IDEATION METRICS

The experiment generated 1,321 total ideas. To thoroughly explore the range of effects of distance and commonness, we applied a range of ideation metrics to these ideas: (1) the extent to which solution features were transferred from examples, (2) quantity of ideation, (3) breadth of search through the space of possible solutions, (4) quality of solution concepts, and (5) novelty of solution concepts. The first three metrics provided measures of the ideation *process* of participants and how they are processing the examples: examining solution transfer provides insight into the mechanisms by which designers might be stimulated by the examples, e.g., did they actually use solution elements; quantity of ideation measures how participants are exploring the design space, e.g., generating/refining a few ideas vs. exploring multiple concepts, which is associated with higher likelihood of generating high-quality concepts [24]; finally, breadth of search provides a measure of the ability to generate a wide variety of ideas, which is associated with the ability to restructure problems, an important component of creative ability [25]. The final two metrics focus on the ideation *products* of participants. We investigate quality because in design, a baseline requirement is that concepts must meet customer specifications. We investigate novelty because there is a high degree of consensus in the literature that creative products are at least novel [26, 27].

4.1 Data pre-processing

The raw output of each participant was in the form of sketches and/or verbal descriptions of concepts. Examples of participant-generated solution concepts are shown in Fig. 2. A number of pre-processing steps were necessary to prepare the data for coding and analysis.

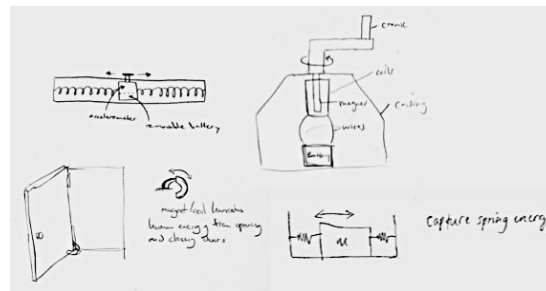


Figure 2: Example participant solution concepts.

First, each participant's raw output was segmented by a trained coder into solution concepts. A sketch and/or verbal description was segmented as one solution concept if it was judged to describe one distinct solution to the design problem. Variations of solutions (e.g., with minor modifications) were counted as distinct solution concepts. Segmentation was independently checked by a second coder. Inter-rater agreement was high (96%), and all disagreements were resolved by discussion. Next, sets of two senior mechanical engineering students rated each solution concept as meeting or not meeting the minimum constraints of the design problem, as described above, to remove off-topic inspirations generated by the patent examples, especially in the second phase. Inter-rater agreement was acceptable, with an average Cohen's kappa of 0.72. All disagreements were resolved through discussion. The 1,066 solution concepts remaining after pre-processing constituted the final data set for analysis.

4.2 Solution transfer

Solution transfer was defined as the degree to which a given participant's idea set contains solution features from the examples s/he received, adjusted for baseline occurrence of those solution features. Key features were generated by one of the co-authors for each of the eight patent examples in the 2 x 2 distance by commonness matrix, and the list was cross-checked for relevance by the other co-authors. A total of 39 key features were identified. Each example was associated with between 4 to 7 features ($M = 4.9$, $SD = 1.0$). Solution concepts were coded for the presence or absence of each of the 39 features. The first 50% of solution concepts was double-coded by two senior mechanical engineering students to establish reliability. Later, all coding was completed by one student only. Test-retest measures of reliability were obtained *in lieu* of inter-rater reliabilities. Cohen's kappa averaged across features was 0.57. Because some features overlapped across examples (e.g., "built into ground, stationary, or permanent" was associated with 4 patent examples), there was not a simple one-to-one mapping of features to examples. To reduce spurious effects from features with low reliability or high

overlap across examples, the initial set of 39 features was filtered down to 23 such that the remaining features: 1) had acceptable inter-rater agreement, i.e., Cohen's kappa greater than 0.4, 2) were not shared by more than three examples, and 3) were not too common, i.e., base rate (collapsed across conditions) less than 0.5. After filtering, the number of features ranged from 1 to 5 ($M = 2.9$, $SD = 1.4$) per example and 4 to 8 ($M = 5.8$, $SD = 1.7$) per each of the four conditions in the distance by commonness 2 x 2 design. Cohen's kappa averaged across filtered features was 0.66.

Solution transfer scores for each participant were computed as follows. First, for each cell in the 2 x 2 (distance x commonness) matrix, we computed for each participant the proportion of his/her ideas that had at least one solution feature from that cell. Next, this proportion was converted into a standardized z-score by subtracting the mean and dividing by the standard deviation of proportion scores for the set of participants *not* in that participant's 2 x 2 cell. This transformation allowed us to separate the probability of participants using solution features from examples they have seen from the probability of using those solution features even if they had never seen the examples. For each participant, the transfer score was the z-score of each feature relevant to the examples they actually received.

4.3 Quantity of ideation

Quantity of ideation was defined as the number of solution concepts generated post analogy, i.e., from the second phase of ideation onwards, that met the minimum constraints of the design problem, viz. (1) the device generates electricity, and (2) it uses human motion as the primary input. As noted in the introduction, quantity is often taken to be a key component of creativity. Quantity was defined at the level of the participant, i.e., each participant received a single quantity score. Because we were primarily interested in the effects of examples on quantity, analyses concentrated on the number of solution concepts generated after receiving examples (i.e., after the first phase) adjusting for the number of solution concepts generated in the first phase (which acted as a covariate to adjust for baseline variation in quantity across participants).

4.4 Breadth of search

Breadth of search was conceptualized in our study as the proportion of space of possible solutions searched by a given participant. To determine the space of possible solutions, the design problem was first functionally decomposed into potential sub-functions by one of the authors, drawing from the reconciled function and flow basis of Hirtz and colleagues [8]. The sub-functions used to determine the space of possible solutions were: 1) import human, 2) transform human energy to mechanical energy, 3) import alternative energy, 4) transform alternative energy to mechanical energy, and 5) transform mechanical energy to electrical energy. Each sub-function solution consisted of a *how* and *what* component, where the former specifies the component of the solution concept that implements the sub-function, and the latter specifies either the input or the output of the sub-function (whichever is the less specified). For example, a solution for the sub-function "import human" might be "*foot with pedals.*"

Two senior mechanical engineering students independently coded the solutions to the sub-functions for each solution concept. The solution types for the *how* and *what* components of each sub-function were generated bottom-up by the students as they coded, with each new solution type being added to a running list of solution types; the running list of solution types for each sub-function constituted the coding scheme. All sub-functions occurred often enough for stable estimates of breadth and novelty (i.e., base rate greater than 0.1, collapsed across conditions). Inter-rater reliability was high, with an average Cohen's kappa across sub-functions of 0.84. All disagreements were resolved by discussion.

We defined the space of possible solutions for each of the *what* and *how* components of each sub-function by enumerating the number of distinct solution types generated by participants across all phases of ideation. A breadth score b_j for each participant on sub-function j was then computed with Eq. (1):

$$b_j = \sum_{k=1}^n w_k \times \frac{C_{jk}}{T_{jk}} \quad (1)$$

where C_{jk} is the total number of solution *types* generated by the participant for level k of sub-function j , T_{jk} is the total number of solution *types* produced by *all* participants for level k of sub-function j , and w_k is the weight assigned level k . To give priority to breadth of search in the *what* space (types of

energy/material manipulated), we gave a weight of 0.66 to the *what* level (which was assigned to $k = 1$), and a weight of .33 to the *how* level (which was assigned to $k = 2$). An overall breadth score for each participant was given by the average of breadth scores for each of the three sub-functions j .

4.5 Quality

Quality of solution concepts was measured using holistic ratings on a set of sub-dimensions of quality. Two other senior mechanical engineering students independently coded solution concepts on 5-point scales ranging from 0 to 4 (0 is unacceptable and 4 is excellent) for six sub-dimensions of quality, corresponding to a set of possible customer specifications: 1) cost, 2) feasibility of materials/cost/manufacturing 3) feasibility of energy input/output ratio, 4) number of people required to operate device at a given moment, 5) estimated energy output, 6) portability, and 7) time to set up and build, assuming all parts already available at hand.

These sub-dimensions were generated by the second author, who is a Ph.D. candidate in mechanical engineering focusing on design methods and cognition, and checked for validity by two other authors, who are mechanical engineering faculty specializing in engineering design. For each sub-dimension, each point on the 5-point scale was anchored with a unique descriptor. For example, for the “feasibility of energy input/output ratio” sub-dimension, 0 was “unfeasible design or input energy completely dwarfs output”, 1 was “input less than output”, 2 was “I/O about even”, 3 was “sustainable/little surplus output; human input easy”, and 4 was “output significantly higher than input.”. The average inter-rater correlation across sub-dimensions was 0.65, and the range was from 0.49 to 0.77. An overall quality score was computed for each solution concept, as given by Eq. (2):

$$Q = \frac{\sum_{j=1}^n q_j \times r_j}{Q_{max}} \quad (2)$$

where q_j is the quality score for quality sub-dimension j , r_j is the reliability of the coding for that sub-dimension, and Q_{max} is the maximum possible overall quality score, which would be given by setting q_j to 4 for each sub-dimension. Since the overall quality score is essentially a proportion of the maximum possible quality score, the score ranges from 0 to 1. The rationale for the weighting by reliability was that we wanted the quality score to be as precise as possible. Inter-rater agreement at the level of this composite score was acceptable ($r = 0.68$).

4.6 Novelty

Novelty was defined as the degree to which a particular solution type was unusual within a space of possible solutions. Recall that for the breadth metric, the space of possible solutions was defined in terms of a set of five core sub-functions for the design problem, with each sub-function decomposed further into *what* and *how* components. Rather than computing novelty scores for solutions to each level of each sub-function (the *what* and *how* levels), we chose to compute novelty scores for the *conjunction* of *what* and *how* solution components for each sub-function. For example, rather than computing the relative unusualness of the solution components “foot” and “pedals” separately for the solution “foot with pedals” for the sub-function “import human interaction,” the relative unusualness of the solution “foot with pedals” relative to other solutions would be computed. The rationale for this choice was that these words in conjunction as a solution have a specific meaning that needed to be considered. Novelty scores were computed for each sub-function solution using Eq. (3), which is a formula adapted from [25]:

$$N_i = \frac{T_i - C_i}{T_i} \quad (3)$$

where T_i is the total number of solution *tokens* generated for sub-function i in the first phase of ideation (collapsed across all participants), and C_i is the total number of solution tokens of the current solution *type* in the first phase of ideation. Because this measure is essentially a measure of proportion, the novelty score for each idea ranges from 0 to 1, with 0 representing solution types found in every solution (this extreme was never observed) and 1 representing solution types that never occurred in the first phase. The initial set of solution concepts (generated in the first phase of ideation) was taken to be the original design space of the participants since it corresponded to concepts generated prior to

receiving examples. The final novelty score for each solution concept was the average of its sub-function novelty scores.

5 RESULTS

5.1 Relationships between metrics

Analysis of the inter-correlations between the ideation metrics suggested a preliminary process model useful for conceptually organizing the results; of course, correlations *per se* do not guarantee causation and other causal models are possible.

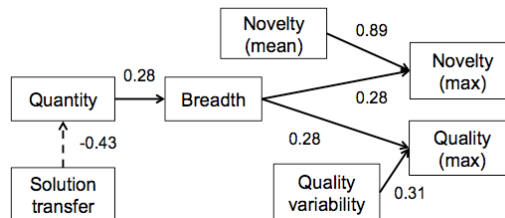


Figure 3: Summary of inter-metric correlations. Numbers shown are Pearson's r . All correlations are significant at $p < 0.01$ ¹.

In this model, increased solution transfer results in decreased quantity, possibly because it becomes difficult to think of solutions beyond the ones presented. Also, a high quantity of ideation allows for greater breadth of search, even if only on a statistical sampling basis. Greater breadth of search, perhaps also only on a statistical sampling basis, in turn allows for generating more novel and higher quality concepts. Additionally, repeatedly searching on the fringes of the design space (as measured by high average novelty) further increases the probability of finding a highly novel concept. Finally, increasing the variability of the quality of solution concepts increases the probability of generating a high quality concept. This finding is similar to of Ulrich and colleagues in the field of innovation management, who have shown empirically that one way to increase the likelihood of finding high market potential product concepts is to increase the variance of the quality of generated concepts [23].

5.2 Effects of example analogical distance

Separate 3-way (distance x commonness x modality) analysis of variance (ANOVA) models were computed for each process variable in the model. In some cases (indicated in each case), the level of that variable during the pre-analogy phase was used as a covariate in the analysis because the baseline measure was a significant predictor of post-analogy performance.

First, there was a main effect of distance on *solution transfer* ($p < 0.01$), with solution elements from far-field examples being much more likely to be used than near-field example solution features ($d = 0.60$ ²; see Fig. 4, bottom left). Next, there was a main effect on *quantity* ($p < 0.01$): receiving far-field examples resulted in significantly fewer solution concepts relative to near-field examples ($d = -0.30$; see Fig. 4, upper left). There were no significant differences in quantity between control (no examples) and either far-field or near-field examples. The small effect of distance on quantity did not translate into an effect on *breadth*: there were no reliable effects of distance on breadth of search ($p = 0.78$).

There were no effects of on either mean or maximum *quality*. However, there was a main effect on the variability in quality of solution concepts ($p < 0.05$; see Fig. 4, lower right): far-field examples resulted in a larger standard deviation in quality of solution concepts relative to either near-field examples ($d = 0.64$) or no examples ($d = 0.78$). There were no significant differences between near-field examples vs. no examples. Finally, there was a main effect on mean *novelty* ($p < 0.05$), where far-field examples resulted in solution concepts that were more novel on average than with near-field examples ($d = 0.56$; see Fig. 4, upper right). Distance similarly impacted maximum novelty of solution

¹ p values denote the probability of observing the data by chance (i.e., if there actually was not a real effect); p values less than 0.05 are conventionally considered statistically significant

² d statistics estimate the size of the difference in group means in terms of the average standard deviation of the two groups in the contrast; in this case, $d = 0.60$ estimates that the mean probability of transfer is greater with far-field vs. near-field examples by 0.60 of a standard deviation (a moderate to large difference)

concepts ($p < 0.05$), with the most novel concept generated with far-field examples being more novel than the most novel concept generated with near-field examples ($d = 0.56$). There were no significant differences in terms of either mean or maximum novelty between control and near-field or far-field examples.

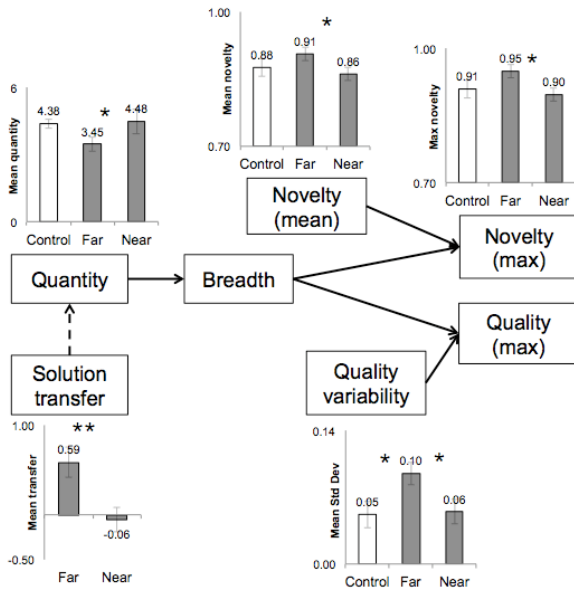


Figure 4: Summary of effects of example distance. *, $p < 0.05$, **, $p < 0.01$. Control group data are shown in white bars. Error bars are ± 1 standard error.

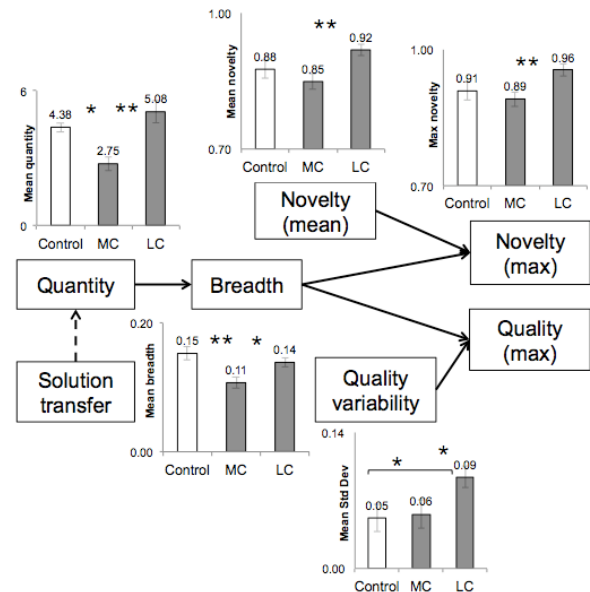


Figure 5: Summary of effects of example commonness. *, $p < 0.05$, **, $p < 0.01$. Control group data are shown in white bars. Error bars are ± 1 standard error.

5.3 Effects of example commonness

Turning now to the main effects of commonness in the same ANOVAs, there were no reliable effects of commonness on *solution transfer* ($p = 0.30$). However, there was a main effect on *quantity* ($p < 0.01$), with more-common examples resulting in significantly fewer concepts vs. less-common examples ($d = -0.67$) or no examples ($d = -0.76$; Fig. 5, upper left). There were no significant differences in quantity between less-common vs. no examples (control). There was also a main effect on *breadth* of search ($p < 0.01$), with more-common examples resulting in less search of the design space vs. either less-common ($d = -0.61$; Fig. 5, lower middle) or no examples ($d = -1.03$). There were no significant differences in breadth between less-common vs. no examples (control).

Similar to example distance, there were no reliable effects of commonness on mean or max *quality*. However, there was a main effect of on variability in quality of participants' solution concepts ($p < 0.05$, $\eta^2 = .06$; see Fig. 5, lower right), with less-common examples resulting in a larger standard deviation in quality of solution concepts vs. either more-common ($d = 0.62$) or no examples ($d = 0.68$). There were no significant differences between receiving more-common examples vs. no examples. Finally, again similar to example distance, there was a main effect of commonness on mean *novelty* ($p < 0.01$), with less-common examples resulting in higher average novelty vs. more-common examples ($d = 0.61$; see Fig. 5 upper right). There was also a main effect on maximum novelty ($p < 0.01$), where the most novel concept generated with less-common examples was more novel on average than the most novel solution concept with more-common examples ($d = 0.61$). There were no significant differences between no examples (control) vs. more- or less-common examples on either mean or maximum novelty.

5.4 Joint effects of example distance and commonness on novelty

While far-field and less-common examples separately increased novelty of ideas, neither far-field examples as a whole nor less-common examples as a whole were significantly different from control, which sat in the middle. To examine whether the *combination* of far-field and less-common properties increased novelty over control, we used a Dunnett's multiple comparison *post-hoc* test. Since there

were no effects of modality on novelty (described below), we collapsed across the picture and text factors and conducted the *post-hoc* test comparing each of the combinations in the 2 x 2 matrix (distance x commonness) with the control condition as a reference group. The *post hoc* test showed that the combination of far-field, less-common examples did in fact increase novelty vs. control, for both mean ($d = 1.14$; see Fig. 6) and max ($d = 1.29$).

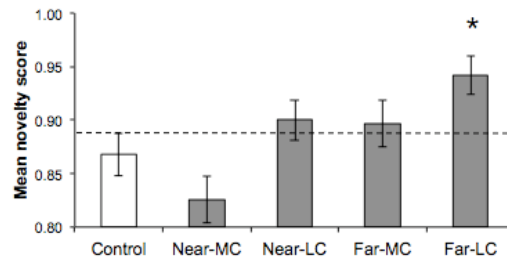


Figure 6. Mean novelty of solution concepts by example distance and commonness. *, $p < 0.05$. Error bars are ± 1 standard error.

5.5 Effects of example modality

Turning to the effects of modality, there was a main effect of on *solution transfer* ($p < 0.01$), with more transfer from text vs. picture examples, regardless of distance or commonness ($d = 0.60$). There was also a main effect on *quantity* ($p < 0.01$), with text examples resulting in significantly fewer concepts than with picture ($d = -0.67$) or no examples (control; $d = -0.56$). There were no significant differences between picture examples and control. There were no additional effects of modality on the other dependent measures (breadth, $p = 0.11$; mean novelty, $p = 0.20$; max novelty, $p = 0.49$; quality variability, $p = 0.44$). In summary, in contrast to example commonness and distance, modality had little impact on the key end-state outputs of the ideation process.

6 DISCUSSION

6.1 Summary of findings

Our findings demonstrate that key dimensions of examples can influence their impact on designers' ideation in important ways. First, augmenting ideation with far-field examples brings significant benefits vis-à-vis the kinds of concepts that can be generated; specifically, ideation with far-field examples enhances the ability to generate highly novel solution concepts and also allows for more variability in the quality of concepts, which may increase the likelihood of generating high quality concepts. However, far-field examples also reduced overall quantity of ideation relative to near-field or no examples. One potential explanation for this is that the inspiration from far-field examples can come with an initial processing cost. With a 3-way ANOVA model on quantity for only the final phase of ideation, removing from consideration quantity of ideation while processing examples, the effects of distance were no longer present ($p = 0.47$). This suggests that the reduction in quantity comes from extra time taken to map the far-field examples to the design problem. Second, the use of less-common examples can positively impact ideation, e.g., as in our findings, increased quantity of ideation, breadth of search, and higher novelty of ideas relative to more-common examples. In contrast to the effects of distance, these beneficial effects do not appear to come with an initial processing cost. Importantly, representation modality of examples did not change the effects of the distance and commonness factors on ideation. However, there was evidence that text representations decreased quantity of ideation; similar to the effects of distance, this suppression effect of text representations can be interpreted in terms of initial processing costs: in the final phase of ideation, the effect was smaller (text vs. pictures $d = -0.32$) and no longer statistically significant ($p = 0.07$). As an ancient proverb puts it, one picture may be worth 1,000 words with respect to conveying design concepts.

6.3 Caveats

The current work comes with a number of caveats. First, we examined the effects of particular examples rather than a range of examples sampled multiple times from a class of examples. This experimental design choice made it more feasible to analyze solution transfer, but raises possibilities of effects being caused by odd examples or example descriptions. To reduce this threat, we had two

examples per condition, and the factorial design of the study permits for multiple replications of main effects. Second, our participants were senior-level engineering students, for the most part, rather than expert designers, and there is some research to suggest that novices have more difficulty with analogical mappings [15, 28]. Finally, our study focused only on the earliest ideation phase, and future work will have to examine the effects of examples on downstream, and in particular finished, solutions. This restriction was most salient in the analyses of quality in that many of the ideas were not feasible or not fleshed out sufficiently to determine feasibility. However, a number of studies point to early ideation as a key moment for intervention to generate innovative designs [29, 30].

6.4 Practical implications and future work

Beyond showing that certain kinds of examples are better than others, our data provide evidence that design-by analogy can confer benefits over and above ideating without examples; specifically, if the goal of conceptual ideation is to ultimately generate and develop a concept that is high quality *and* novel, then using analogies (specifically far-field, less-common examples) is worth the extra effort over not using analogies.

There are also implications for the design of tools and methods to support design-by-analogy. Given the demonstrated benefits of far-field analogies, and taking into account humans' difficulty in retrieving these analogies from memory [14], computational tools that are able to define and compute functional and surface similarity between items in a design space in a principled manner relative to the current design problem would hold excellent potential as aids for inspiration. These tools could maximize the potential benefits of these analogies by retrieving and delivering them to designers in a timely manner. Additionally, if these systems are able to give priority to analogies that are relatively unusual or infrequently encountered, the potential for inspiration might be even higher. Currently, the state of the art for computational design-by-analogy tools has not reached the point of being able to provide flexible and real-time support in this manner. The present work provides an impetus for investment into this important research area, as the potential benefits to engineering practice and to society via increased innovation is high.

Finally, there are interesting connections to some other work on designers' preferences with respect to analogical distance in choosing analogies. Recent work has shown that designers' have a preference for finding inspiration from near-field rather than from far-field analogies, and that this preference can be optimal with respect to utility of inspired designs, as rated by end users [31]. Taken together with the findings reported in this paper on the effects of far-field analogies, one potential implication is that a nuanced approach with respect to analogical distance may more effectively maximize the benefits of analogies; specifically, rather than only and always focusing on far-field analogies, it might be more optimal to draw from far-field analogies primarily during the early stages of conceptual design, where broad search and novelty are at a premium, and then subsequently switch to near-field analogies to help narrow the search space to more viable solutions. Of course, this implication is primarily speculative at this point, and more evidence is needed to verify if such a strategy would in fact be effective; future work should explore this intriguing and potentially impactful line of research.

Overall, the present work has helped to show the benefits of analogies for innovative design, and perhaps more importantly, suggested important guidelines for maximizing the benefits of design-by-analogy while avoiding potential pitfalls, and promising further lines of research to increase knowledge on effective design-by-analogy practices.

ACKNOWLEDGEMENTS

This work is supported by grants from the National Science Foundation, grant numbers CMMI-0855326, CMMI-0855293, and CMMI-0855510.

REFERENCES

- [1] P. Leclercq, P., and A. Heylighen, 2002, "5,8 analogies per hour", *Artificial Intelligence in Design '02*. J. S. Gero (ed.), Cambridge, UK, pp. 285-303.
- [2] B. T., Christensen and C. D. Schunn, 2007, "The relationship of analogical distance to analogical function and pre-inventive structure: The case of engineering design. *Mem. & Cog.*, **35**(1), pp. 29-38.
- [3] J. Linsey, J. Murphy, J. Laux, A. Markman, and K. L. Wood, 2009, "Supporting Innovation by Promoting Analogical Reasoning," *Tools of Innovation*, Oxford University Press.

- [4] D. Gentner, 1983, "Structure-Mapping: A Theoretical Framework for Analogy," *Cog. Sci.*, 7, pp. 155-170.
- [5] W. J. J. Gordon, 1961, *Synectics: The development of creative capacity*, Harper and Brothers, New York.
- [6] M. French, 1988, *Invention and evolution: Design in nature and engineering*, Cambridge University Press, Cambridge, UK.
- [7] E. Hacco and L. H. Shu, 2002, "Biomimetic concept generation applied to design for remanufacture," *2002 ASME Des. Eng. Tech. Conf. and Comp. and Info. in Eng. Conf.*, Montreal, Quebec, Canada.
- [8] J. Linsey, K. L. Wood, and A. Markman, 2008, "Increasing Innovation: Presentation and Evaluation of the WordTree Design-by-Analogy Method," *ASME International Design Theory and Methodology Conference*, Brooklyn, New York, DETC2008-49317.
- [9] J. Hirtz, R. B. Stone, and D. A. McAdams, 2002, "A functional basis for engineering design: Reconciling and evolving previous efforts," *Res. Eng. Des.*, **13**, pp. 65-82.
- [10] S. M. Smith, T. B. Ward, and J. S. Schumacher, 1993, "Constraining effects of examples on a creative generation task," *Mem. & Cog.*, **21**(6), pp. 837-845.
- [11] D. G. Jansson and S. M. Smith, 1991, "Design fixation," *Des. Stud.*, **12**(1), pp. 3-11.
- [12] D. W. Dahl and P. Moreau, 2002, "The influence and value of analogical thinking during new product ideation," *J. Mark. Res.*, **39**(1), pp. 47-60.
- [13] J. O. Wilson, D. Rosen, B. A. Nelson, and J. Yen, 2010, "The effects of biological examples in idea generation," *Des. Stud.*, **31**, pp. 169-186.
- [14] K. D., Forbus, D. Gentner, D., & K. Law, 1994, "MAC/FAC: A model of similarity-based retrieval," *Cog Sci*, **19**, pp. 141-205.
- [15] H. Casakin and G. Goldschmidt, 1999, "Expertise and the use of visual analogy: Implications for design education," *Des. Stud.*, **20**(2), pp. 153-175.
- [16] R. W. Weisberg, 2009, "On "out-of-the-box" thinking in creativity," In A. B. Markman & K. L. Wood (Eds.), *Tools for Innovation*. New York: Oxford University Press.
- [17] K. Duncker, 1945, *On Problem Solving*. Washington, DC: Amer. Psych. Assoc.
- [18] R. E. Adamson, 1952, "Functional fixedness as related to problem solving: A repetition of three experiments," *J Exp. Psych.*, **44**(4), pp. 288-291.
- [19] N. R. F. Maier, 1931, "Reasoning in humans. II. The solution of a problem and its appearance in consciousness," *J. Comparative Psych.*, **12**, pp. 181-194.
- [20] W. G. Chase and H. A. Simon, 1973, "The mind's eye in chess," In W. G. Chase (Ed.), *Visual Information Processing*. New York: Academic Press.
- [21] M. T. H. Chi, P. J. Feltovich, and R. Glaser, 1981, "Categorization and representation of physics problems by experts and novices," *Cog. Sci.*, **5**, pp. 121-152.
- [22] S. Ohlsson, 1992, "Information-processing explanations of insight and related phenomena," In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking: Vol. 1*. Hertfordshire, UK: Harvester Wheatsheaf.
- [23] I. Tseng, J. Moss, J. Cagan, and K. Kotovsky, 2008, "The role of timing and analogical similarity in the stimulation of idea generation in design," *Des. Stud.*, **29**, pp. 203-221.
- [24] K. Girotra, C. Terwiesch, and K. T. Ulrich, 2010, "Idea generation and the quality of the best idea," *Management Sci.*, **56**(4), pp. 591-605.
- [25] J. J. Shah, N. Vargas-Hernandez, and S. M. Smith, 2003, "Metrics for measuring ideation effectiveness," *Des. Stud.*, **24**, pp. 111-134.
- [26] A. B. Markman and K. L. Wood (Eds.), 2009, *Tools for innovation: The science behind practical methods that drive new ideas*, Oxford University Press, New York, NY.
- [27] M. A. Boden, 2004, *The Creative Mind: Myths and Mechanisms (2nd Ed.)*, Routledge, London.
- [28] L. R. Novick, 1988, "Analogical transfer, problem similarity, and expertise," *J. Exp. Psych: LMC*, **14**(3), pp. 510-520.
- [29] C.M. Vogel, J. Cagan, and P. Boatwright, 2005, *The Design of Things to Come*, Wharton School Pub.
- [30] C. Terwiesch and K. T. Ulrich, 2009, *Innovation Tournaments*, Harvard Business School Pub.
- [31] B. T. Christensen, 2010, "Images of users and products shown during product design increase users' willingness-to-use the innovation," *AIEDAM*, **24**, pp 179-189.