Fixation or inspiration? A meta-analytic review of the role of examples on design processes

Ut Na Sio and *Kenneth Kotovsky*, Department of Psychology,
Carnegie Mellon University, Pittsburgh, PA 15213, USA
Jonathan Cagan, Department of Mechanical Engineering,
Carnegie Mellon University, USA

*A meta-analytical review of design studies (*N = *43) was conducted examining whether and under what conditions the presence of examples will induce fixation or inspiration. The analysis revealed that providing examples made individuals generate more example-related and fewer categories of ideas, however, the ideas produced were more novel. Also, the quality of solutions ideas was positively correlated with the degree of copying from examples. The facilitatory effects on novelty and quality increased when fewer and less common examples were presented. Presenting a single and uncommon example may encourage individuals to shift from traversing between different parts of the problem space to conducting a deeper search in a specific and remote domain, facilitating the generation of high-quality and novel ideas.*

Keywords: fixation, creativity, innovation, design process(es), design strategy

In design, people often seek examples for inspiration. Design-by-analogy, using examples to spark ideas on what a solution to a problem might be, is a common approach to innovation concept generation (Goel, 1997; Hey, Linsey, Agogino, & Wood, 2008; Holyoak & Thagard, 1995). Even though example solutions may not always provide fully developed solutions to design problems, they can serve as cues that help retrieve relevant concepts from long-term memory and in turn facilitate the development of innovative conceptual designs. However, one risk with this approach is that example solutions, even potentially beneficial ones that do not violate design constraints, may negatively impact design processes in the form of design fixation, directing individuals to search in the example-related domains only (Chan et al., 2011; Jansson & Smith, 1991). Do examples induce fixation or inspiration? Findings of the previous studies investigating the effects of examples on design processes are inconclusive, with some reporting the fixation effect induced by the presence of examples (Chrysikou & Weisberg, 2005; Jansson & Smith, 1991); and others demonstrating how the presence of examples inspires the design idea generation processes (Helms, Vattam, & Goel, 2009). The co-existence of the positive and negative effects elicited by examples

Corresponding author:
Ut Na Sio
unsio@cmu.edu

www.elsevier.com/locate/destud
0142-694X *Design Studies* **39** (2015) 70–99
http://dx.doi.org/10.1016/j.destud.2015.04.004

makes it difficult to judge whether exposure to examples is helpful or not. A comprehensive theory or model that can explain how and why examples might impact the design processes is still absent. To address these, we conducted a meta-analytical review of the past studies in this area, in an attempt to assess the impacts of examples on design processes and to identify under what conditions the presence of examples will induce fixation or inspiration. Understanding the impact of examples on design processes allows us to make an effective use of examples to promote innovation and creativity in design practice. Further, examining how the external information (e.g., examples) influences individuals' design processes also enriches our understanding of problem solving in which the detailed interaction between the external information and the individual problem solving processing is still poorly understood.

1 Fixation or inspiration

When designing a new product, individuals usually do not create something new from scratch. Instead, they transform, combine, or adapt elements of existing designs to generate new ideas (Ward, 2007). However, looking at existing designs may not always provide inspiration to individuals. Empirical studies on design processes have revealed that consulting existing designs may instead negatively impact the quality of the design solution. When individuals are given an example solution to look at, they often tend to produce a solution similar to the example provided. Individuals do not copy ideas only from relevant examples that fulfill the task requirement, design replication occurs even when the example is a poor one that does not fit the task requirement. Jansson and Smith (1991) observed and described a phenomenon called *design fixation* in engineering design that engineering design students tended to incorporate features of the examples in their design ideas even when the features of the examples were negative features. Studies have been conducted to examine how design fixation can be overcome by manipulating the properties of the examples, e.g., common-ness of the examples (Purcell & Gero, 1991), richness of the examples (Cardoso & Badke-Schaub, 2011), and modality of the examples (Chan et al., 2011; Viswanathan & Linsey, 2013a, b).

In contrast to the studies examining design fixation, there is another line of research examining how examples serve as analogies and facilitate creativity. Anecdotal reports of creative discoveries and inventions have supported the potential of exemplar analogies for creativity, such as 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. Empirical works have been reported showing that presenting examples to individuals can inspire creativity (Casakin & Goldschmidt, 1999; Goel, 1997). The examples used in the design-by-analogy studies are potentially beneficial examples that do not violate the task requirement (termed "non-negative" examples in this paper), and are distant or uncommon examples. For example, biological-inspired engineering design uses analogies of biological systems to develop solutions for engineering problems (Helms et al., 2009). Also, different procedural factors were manipulated in these studies, e.g., number of examples and timing of presentation (before vs. during problem solving), in an attempt to identify the optimal settings to facilitate analogical transfer from the examples to the problem.

This paper will review and integrate at a meta-level the findings of these studies, which vary in terms of the properties of examples as well as the experimental settings. This should allow us to find out the overall impact of examples on design processes and more importantly identify the factors that can moderate the magnitude of the exemplar effects. In this meta-analysis, we focus on studies using non-negative examples, examining the effects of these examples as to whether they induce design fixation or facilitate the generation of ideas.

Examining the impact of non-negative and negative examples are both crucial in understanding how to improve design pedagogy and to enhance design creativity. Studies using negative examples have been conducted with a focus on examining how individuals overcome fixation induced by misleading concepts and ideas. But our paper focuses on another way to enhance design creativity, examining when and how individuals make use of examples that can potentially facilitate design processes. Studies using negative examples are excluded because of the specific objective of this meta-analytic review. In the following section, we will first discuss the cognitive processes underlying design problem solving before going into the details of how non-negative examples can potentially impact the design processes.

1.1 Design as search

Designing can be viewed as a form of complex problem solving, and the cognitive model of problem solving can also be used to describe the design processes (Cagan, Kotovsky, & Simon, 2001; Goel, 1995; Simon, 1969). Similar to problem solving, design can be decomposed into multiple stages: when individuals are given a design problem to solve, the first step is to encode the task information and create a problem space consisting of task-relevant knowledge retrieved from long-term memory; designing is then searching for task-relevant concepts within the problem space, and elaborating and integrating these concepts to generate a design solution that meets the design goals. Usually, the design goals and the requirements for good solution are not fully specified in advance and there is more than one possible approach to solve the problem. Thus, the problem space can be so large that it will be difficult to search efficiently. To deal with this, individuals usually self-impose additional task constraints to frame the problem more specifically and confine the search to certain parts of the problem space. If the search yields no solution that meets the design goals, individuals may then have to expand the search by relaxing or surmounting the constraints.

The characteristics of the design solutions are determined by how individuals explore the problem space. The types of ideas individuals generate depend on where individuals search for solution ideas. For example, individuals who are fixated on the example-related domain tend to generate more example-related ideas (Jansson & Smith, 1991), and individuals who search for ideas in a less-common domain are more likely to produce novel ideas (Purcell & Gero, 1996). Also, individuals often do not just focus on one domain, instead they sample ideas from different domains, and the more domains they have explored the higher the variety as well as quantity of ideas that will be generated. A deep exploration of the problem space should lead individuals to generate high novelty and quality ideas (Rietzschel, Nijstad, & Stroebe, 2007). Examining the characteristics of a design solution should offer us a window into the search strategy an individual adopted when solving a design task.

1.2 Design solution evaluation

Among studies examining the effects of examples, design solutions were usually assessed using at least one of the following metrics: degree of copying, quantity, novelty, variety, and quality of the ideas embedded in the design solution (Shah, Smith, & Vargas-Hernandez, 2003). These measures are good indicators of the breadth and the depth of the solution search. *Degree of copying* is the number of example-related ideas embedded in the design solution. Duplicating the exemplar features is taken as an indication that individuals are confined to search within the example-related space. *Quantity* is the total number of ideas generated. *Novelty* is a measure of how unusual or unexpected an idea is as compared to other ideas. In design studies, the novelty of an idea is judged in relation to how uncommon it is in the overall population of ideas. *Variety* is a measure of the explored solution space during the idea generation process. *Quality* is a measure of the feasibility of an idea. Examining the effect of examples on degree of ideas, the quantity, quality, variety and novelty should allow us to find out how the search strategy is impacted by the presence of examples.

Early design studies examining the effects of examples have reported that individuals tend to copy the ideas embedded in the examples and fixate on example-related domains when searching for solution ideas (Jansson & Smith, 1991). This may narrow down the scope of the solution search and in turn reduce the quantity and the variety of the solution ideas. However, a focused and narrow search may facilitate a deeper exploration in the problem space, and this could increase the likelihood of generating high quality and novel design solutions. Therefore, a larger degree of copying may be associated with a narrower but deeper search, and this may not be harmful, and even beneficial, to design processes. In this study, we conducted a systematic meta-analytic review using statistical analysis to evaluate the overall effect of examples on the degree of copying, and the quantity, novelty, variety, and quality of the solution ideas, and to identify the moderators that modify the size and the direction of the effects.

1.3 Potential moderators

A variety of factors have been manipulated to examine how they impact the exemplar effects. In this study, we focused on three of these factors that are shown to be impactful in problem solving: the timing of presentation, the common-ness of the example, and the number of examples.

1.3.1 Common-ness

Studies in problem-solving reported that individuals tend to fixate on typical knowledge when solving a problem and this inhibits exploration in the less typical knowledge domain (Duncker, 1945). Presenting a familiar example may further activate the common knowledge, making individuals more fixated in the typical area. Conversely, presenting uncommon or remotely-associated information may help retrieve knowledge in unexplored domains, and this should facilitate novel conceptual integration. Empirical work has also demonstrated that looking at examples that are remotely associated with the design tasks was less likely to induce fixation (Purcell & Gero, 1992) and was more likely to lead to the generation of innovation solution (Chan et al., 2011; Chiu & Shu, 2012; Dahl & Moreau, 2002).

However, Fu and Chan et al., (2013) and Fu and Murphy et al., (2013) revealed that consulting uncommon examples may not be helpful all the time, especially when the examples are too far from the problem and have no obvious connection to the problem. Chan, Dow, and Schunn (2014) reported that conceptually close rather than far ideas appeared to be more beneficial for creative production. Should we look at common or uncommon examples for inspiration? It is difficult to draw any cross-study conclusions on this issue because different common-ness measures were adopted among the past studies. For studies comparing the effect of common vs. uncommon examples on design solutions, they usually prepared a set of examples and classified them into two groups: uncommon and common, based on either results of the subjective evaluations (Chan et al., 2011; Purcell, Williams, Gero, & Colbron, 1993) or quantitative analyses, such as the semantic closeness between the problem statement and the description of the examples (Fu & Chan et al., 2013). All these classifications are relative measures of common-ness because they depend on the particular selection of examples included in the same study. Fu and Chan et al., (2013) pointed out that some examples that were considered to be distant examples to a problem in a study were classified as the close ones to the same problem in another study because of the selection of examples for comparison.

In order to interpret and integrate the findings on the common-ness of examples and exemplar effects, re-assessing the common-ness of the example solutions presented in the past studies with a unified measure is needed. In this study, we asked one faculty member and two senior undergraduates in Mechanical Engineering to evaluate the common-ness of the example solutions used in the past studies. They were asked to focus on two criteria: whether the concepts mentioned in the example solution are common knowledge in Mechanical Engineering and whether the example solution is a common approach to that specific problem. The inclusion of both the task-general and task-specific criteria should give us a comprehensive measure on the common-ness of the example solutions. In this meta-analysis, we examined if the common-ness of example solutions can modify the size and the direction of the exemplar effects on the degree of copying, quantity, novelty, variety, and quality of the solution ideas.

1.3.2 Timing

In problem solving research, the findings on the effects of cues are equivocal in that individuals are not always able to make use of cues to solve the problem (Dodds, Smith, & Ward, 2002; Sio & Ormerod, 2009, 2014). Whether solvers have reached an impasse or not may affect their susceptibility or receptiveness to various types of examples (Moss, Kotovsky, & Cagan, 2007, 2011). Moss et al. (2011) demonstrated that cues were more effective when they were presented after a period of initial work on the problem than at the beginning of problem solving. They suggested that after a period of initial work, individuals are more likely to reach an impasse, and therefore, they are more capable in making use of the cues.

Another explanation to account for this timing effect is that after working on a problem for a while, the initial problem space has been expanded and this should shorten the distance between the cues and the current problem representation, making it easier to integrate the cues with the problem. In the area of engineering design, although past studies have shown that far examples are more likely to lead to the generation of innovative solutions, they can be unhelpful or even harmful when they are too distant to the problem (Fu & Chan et al., 2013). Identifying the optimal distance between the example and the problem is critical in making use of examples effectively. However, this distance is not absolute, depending on individuals' current problem presentation. Tseng, Moss, Cagan, and Kotovsky (2008) found that near examples were more impactful than far examples when they were both presented at the beginning of problem solving, and examples that are distantly related to the design problem were more impactful when presented after a period of initial work than before problem solving. According to this explanation, postponing the presentation time of examples should also magnify the exemplar effects, especially when the examples are uncommon.

However, presenting an example after a period of work has begun may not be that impactful due to the sunk cost effect (Arkes & Blumer, 1985) that individuals have a greater tendency to continue an endeavor once an investment in money, effort, or time has been made. Viswanathan and Linsey (2013b) revealed that building physical models can cause a higher degree of design fixation because of the high cost (e.g., time, effort, and money) associated with physical modeling. According to the sunk cost effect, the longer an individual works on the problem the less likely an individual will change their problem-solving approach. It therefore should be better to present the examples at the beginning than during problem solving. In this meta-analysis, we examined whether presenting the examples before or during problem solving would generate a larger positive impact on design processes.

1.3.3 Number of the examples

Recent studies examining the effect of examples on design processes tend to present multiple rather than single example solutions to participants. It is suggested that the example solutions may serve as cues activating concepts that are relevant to the problem and in turn facilitate the design processes. Presenting multiple examples may activate different sets of concepts, and therefore, individuals receiving multiple examples may be more likely to combine different ideas to generate innovative solutions. Also, each single example may lead to different search directions, encouraging a multi-domain solution search.

However, searching requires conscious attention and effort, and humans possess limited attentional cognitive resources making it difficult to search broadly and deeply at the same time. Also, designing is not a pure idea generative process; it also involves conscious evaluation and elaboration of ideas once they are being generated (Williams & Sternberg, 1988). Looking at multiple examples may inhibit individuals from searching deeply, and evaluating and elaborating each idea carefully. Together, these may negatively impact the solution quality. In line with this, a recent study of Kazakci, Gillier, Piat, and Hatchuel (2014) revealed that when solving real-life design tasks, professional design teams that generated fewer initial ideas also generated better final design solutions. McComb, Cagan, and Kotovsky (2015) reported that when solving an open-ended engineering design task, high-performing teams searched a small portion of the design space after a period of initial work, while low-performance teams continued to traverse the design space. These findings converge to suggest a potential link between problem solving performance and search strategies. In this meta-analysis, we compared how the number of examples presented impact the solution search strategy, and in turn, impact the characteristics of the design solutions.

In sum, our statistical meta-analysis addressed the following issues: First, whether there is reliable evidence for the impact of examples on degree of copying, quantity, variety, novelty, and quality of ideas. Second, how the timing of presentation, the common-ness, and the number of examples modify the strength of the exemplar effect on each of these five measures.

2 The meta-analysis

We collected publications that contained studies relevant to the meta-analysis through a search of the ERIC, EBSCOhost, Google Scholar™, MEDLINE, PsycInfo, and PsycArticles databases using the keyword "design", intersected with one of *fixation*, *example*, *analogy*, or *inspiration*. Then, references given in all the obtained articles were systematically searched for additional relevant publications. In total, 16 relevant publications meeting the following criteria were identified and obtained, and were assimilated in the analysis:

1. Examples presented did not violate the task requirements.
2. The study included a control (no example) condition.
3. The study examined individual rather than group design problem solving.
4. The study reported information that allowed the computation of an effect size.

A number of design studies we obtained have examined whether individuals can ignore poor examples containing features that violate the constraints of the task and nonetheless generate acceptable design solutions. However, our interest lies in understanding when and how individuals make use of examples that can potentially facilitate the design processes. Therefore, we only focused on studies using non-negative examples that *did not* violate the task requirements (criterion 1). Note that such examples are not necessarily beneficial in that they may include extremely common features that can induce fixation. The inclusion of criterion 2, a control condition, is essential to provide a baseline against which performance in exemplar conditions can be compared. Studies were not included in the meta-analysis if they did not include a control group. There are some studies examining the effect of examples on group design processes. The quality of the group design solutions depends on not only individuals' problem solving approach but also the quality of the interaction between individuals. The primary interest of this study is to understand how examples impact individuals' search strategy when solving a design task. Therefore, only studies examining individual design problem solving were included in the analysis. The information required for computing effect sizes is discussed in the section "Estimation of effect sizes". Some publications include multiple experiments, thereby allowing a reasonable sample size of independent studies examining the effect of positive examples on design processes ($N = 43$) to be achieved.

2.1 Coding procedure

Many of the experiments reported in the selected publications had two or more example conditions, such as showing pictorial vs. text examples. For the sake of the meta-analysis, experiments with more than one example condition were broken down into independent studies with one example condition and one control condition. The same control group may be included in more than one independent study, and compared with more than one example condition. For example, in Tseng et al.' s experiment (2008), there were one control and three example conditions. The experiment was decomposed into three studies in this analysis. To avoid inflating the degrees of freedom available, the number of participants in the control condition was split across studies entered into the analysis, a method advocated by Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, and van Ijzendoorn (2007).

2.2 Estimate of effect size

For each study entered into the meta-analysis, the effect size, Cohen's *d*, was computed for each measure (degree of copying, quantity, novelty, variety, and quality of ideas). Cohen's *d* in this meta-analysis comprised the difference in the measure between the control and exemplar conditions divided by their pooled standard deviation (Hedges & Olkin, 1985). In some cases, effect sizes had to be calculated from *t*- and *F*-values, frequencies or *p*-values. If a *p*-less-than value was given instead of an exact *p*-value, the *p*-less-than value was treated as an exact value, and an estimate of Cohen's *d* was generated. For studies that did not include any of the above-mentioned information but only provided statements of non-significant differences between the control and the example groups, then Cohen's *d* was assumed to be zero. Among the studies that included multiple example conditions, some provided a statement of non-significant performance differences among the exemplar conditions, and only reported the overall performance difference between the control and the exemplar conditions. In such cases, all exemplar conditions were assumed to generate the same magnitude of effect size.

Following Hedges and Olkin (1985) suggestion for removing bias caused by small sample studies, an unbiased effect size estimate was computed by multiplying the effect size of each single study by a factor $1-3/(4(\text{total N}-2)-1)$, where total N is the total number of participants of that study. Any unbiased effect size larger than 2 standard deviations from the group mean was considered an outlier, and was recoded to the value of the effect size found at 2 standard deviations, following a procedure for reducing the bias caused by extreme effect sizes reported by Lipsey and Wilson (2001).

2.3 Variance in effect sizes between studies

We predicted that the variance in magnitude of the unbiased effect size estimates among studies was not due simply to sampling error but instead to the difference in settings of each study (i.e., the timing, common-ness, and number of examples presented). Regression analyses were conducted to examine the impacts of these moderators on unbiased effect size estimates for degree of copying, quantity, variety, novelty, and quality of ideas.

It is suggested that larger studies include less sampling error, and therefore deserve a larger weight in combining the effect sizes. Hedges and Olkin (1985) revealed that the optimal weight is the inverse of the variance, which is the sum of the within-study variance[1] and between-studies variance (random variance component) of the unbiased effect size. In our regression analyses, each unbiased effect size estimate was weighted by the inverse of its variance.

To determine the random variance component of each unbiased effect size estimate, a heterogeneity test was first carried out to confirm the assumption of heterogeneous distribution of effect sizes for each measure. If a heterogeneity test did not reach significance, then a zero random variance component was assumed. If the test for heterogeneity was significant, then the random variance component was computed.

The standard measure of heterogeneity is the Cochran's Q test. The Q statistic is the weighted sum of squared differences between the unbiased effect size estimate of each independent study and the weighted average unbiased effect size estimate across studies. Q is distributed as a chi-square statistic with $k - 1$ degree of freedom, where k is the number of independent studies. If the Cochran's Q test for heterogeneity is statistically significant (Q is larger than the chi-square value with $k-1$ degree of freedom), the assumption of heterogeneous distribution is supported. The random variance component was equal to $[Q-(k-1)]/c$, where Q is the Cochran's Q value and k is the number of studies. The formula for c was [(the sum of the inverse of the within-study variance) − (the sum of the square of the inverse of the within-study variance)]/(the sum of the inverse of the within-study variance) and was suggested by Cooper and Hedges (1994).

There is a widespread belief that a test of heterogeneity must be found to be significant before analyses can be conducted to examine the impact of moderators on an effect size. However, the heterogeneity test is usually underpowered and often fails to detect the heterogeneity especially when the sample size is small, and a non-significant heterogeneity test result does not preclude the search for moderators (Kontopantelis, Springate, & Reeves, 2013; Rosenthal, 1995). In this study, we conducted weighted least-squares regression analyses to examine the contributions of moderators to the unbiased effect size estimates, regardless of the significance level of the test of heterogeneity.

---

1. The formula for the within-study variance was [(2 × square of total N) × (N of experimental × N of control × square of unbiased effect size)]/(2 × total N × N of experimental × N of control), where N is the number of participants in that condition (Cooper & Hedges, 1994).

2.4 Publication bias

Publication bias arises when the published studies included in the meta-analysis do not represent all studies on the topic of interest. Studies with statistically significant effects are more likely to be published, as compared to those with conclusive results (because of smaller sample sizes or less statistical precision). Prior to investigating the impacts of moderators on effect sizes, a preliminary analysis was undertaken to assess if a publication bias existed in the selection of studies (Macaskill, Walter, & Irwig, 2001). Five weighted least-squares linear regressions were carried out, one for each of the effect size measures, using the unbiased effect size estimates weighted by the inverse of variance as the dependent variable and the sample sizes as the predictor variable. In the absence of any publication bias, there should be no association between effect size and sample size. Therefore, the regression slope (i.e., coefficient of the predictive variable) would be expected to approach zero if there is no publication bias. The outcome of this analysis is reported in the Results section.

2.5 Common-ness

The common-ness of each example solution was assessed in two ways: the common-ness of knowledge and the common-ness of the approach to the problem.[2] The majority of the tasks are functionally-based in that they emphasize the functionality, practically, and human factor attributes of the products. Therefore, one faculty member and two senior students in Mechanical Engineering were asked to look at the problems and the corresponding example solutions, and answer two questions *"Is the example solution a common approach to that design problem*" and "*Are the mechanisms mentioned in the example considered to be common knowledge to mechanical engineering students*", using a 5-point scale. The first question is a task-general measure while the second one is task-specific. The common-ness score of each example was the average of the scores of the two questions. For studies presenting multiple examples to individuals, a single common-ness score, which was the average of the common-ness scores for all the examples, was computed to represent the common-ness of the examples presented.

The common-ness scores assigned by the three raters were moderately correlated with each other. The correlation between the overall common-ness scores given by the two students was $r = .77$, and the correlations between the common-ness scores by faculty member and by each of the students were $r = .75$ and $r = .69$, all $p < .001$. A few studies have their own common-

2. Please see Supplementary Material A for the description of the problem and the corresponding example solutions used in each study. Alternatively, please contact the first author for the materials.

ness measures and the final common-ness score was also correlated with the classification used in these studies, $r = .58$, $p = .04$.

We averaged the scores produced by the three raters to generate an overall common-ness score for each study.

3 Results

43 studies were included in this meta-analysis.[3] The total number of participants was 1229 and the median number of participants per study was 24. Among these participants, 67.1% ($N = 825$) of them were engineering undergraduates, 18.4% ($N = 226$) of them were undergraduates in industrial design or architecture, 14.5% ($N = 178$) of them were undergraduates with unspecified majors. For each study, an unbiased effect size estimate was computed for each of the design solution measures, with zero indicating no difference between the experiment and control group, an effect size of $\pm$ .2 is a small effect, and an effect size of $\pm$ .5 is a medium effect (Cohen, 1988). A standard system was used to code each study. Table 1 presents the coding system.[4]

Preliminary analyses were first conducted to check the presence of publication bias. In all models, the regression coefficients of the predictive variables were not significantly different from zero, all $p > .10$, suggesting a non-significant correlation between sample size and effect size estimates. This supports the absence of publication bias. Thus no correction has been made for publication bias.

3.1 Overall effect of non-negative examples on design process

Analyses were conducted to examine 1) the impact of non-negative examples on the degree of copying, quantity, novelty, variety, and quality of the solution ideas, and 2) whether the proposed moderators can account for the effect size variability among the past studies. To address the first question, the weighted mean of the effect size estimates on each of the five measures was computed. Table 2 presents the random variance component, the weighted mean, standard error, and 95% confidence interval of effect size estimates on each measure.

3. The meta-analysis included these papers: Agogué et al., 2014; Cardoso & Badke-Schaub, 2011; Cardoso, Gonçalves, & Badke-Schaub, 2012; Chan et al., 2011; Dahl & Moreau, 2002; Fu & Chan et al., 2013; Fu & Murphy et al., 2013; Gonçalves, Cardoso, & Badke-Schaub, 2012; Lujun, 2011; Perttula & Sipilä, 2007; Purcell & Gero, 1992; Smith, Ward, & Schumacher, 1993; Tseng, Moss, Cagan, & Kotovsky, 2008; Wilson, Rosen, Nelson, & Yen, 2010; Hassard, Blandford & Cox, 2009.

4. Please see Supplementary Material B for the information extracted from each independent study by using this coding system. Alternatively, please contact the first author for the materials.

Table 1 Coding system

Variable	*Coding description*
Author	Author(s) of the study
Year	Year the study was published
Total	Total number of participants
Background	Fields that participants majored in • Mechanical Engineering, Engineering, and Industrial Design Engineering students were grouped as "Engineering" • *Architecture, Industrial Design and, Interior Design students* were grouped as "Design" • Unspecified
Task	The design task
Example	Example solutions provided
Timing	Timing of the presentation • Before = At the beginning of problem solving • During = After a period of initial work
Numbers of examples	Number of examples presented
Degree of copying	The unbiased effect size estimate on the degree of copying
Quantity	The unbiased effect size estimate on the quantity of solution ideas
Novelty	The unbiased effect size estimate on the novelty of solution ideas
Variety	The unbiased effect size estimate on the variety of solution ideas
Quality	The unbiased effect size estimate on the quality of solution ideas
Common-ness	The overall common-ness score of the examples

As per Table 2, the weighted mean of the unbiased effect size estimates on the *degree of copying* was significantly larger than zero ($M = .46$, $SE = .08$), suggesting that the exemplar group generated more example-related ideas than the control group. Presenting examples to individuals may lead them to spend more time in exploring the example-related domains and generate more example-related ideas. This should also narrow down the scope of search resulting in a reduced variety of solution ideas. In line with this, a negative effect size on the variety of solution ideas was revealed ($M = -.25$, $SE = .11$).

However, the effects of examples were not all negative. The effect size on *quantity* was not significantly different from zero, and its confidence interval ranged from small positive (.19) to small negative (−.11). These results indicated that looking at examples did not significantly reduce the number of ideas

Table 2 The random variance component, the weighted mean, the SE, and 95% confidence interval of the unbiased effect size estimates on each measure

	Degree of copying	*Quantity*	*Novelty*	*Variety*	*Quality*
Number of studies	28	31	29	20	19
Random variance component	.00	.00	.09	.00	.26
Weighted unbiased effect size estimate	.41*	.10	.27*	−.24*	−.02
SE of the weighted unbiased effect size estimate	.09	.08	.10	.11	.15
95% Confidence interval of the weighted mean	.60, .22	.26, −.06	.47, .06	−.02, −.46	.30, −.34

Note. Majority of the studies included in the meta-analysis did not measure the effect of non-negative examples on all these five aspects, and therefore, there were less than 43 studies for each measure.*Significantly different from zero.

Table 3 Correlation between the unbiased effect size estimates for different measures

Variables	*Degree of copying*	*Quantity*	*Novelty*	*Variety*	*Quality*
Degree of copying	1				
Quantity	−.51* (n = 23)	1			
Novelty	.24 (n = 14)	.16 (n = 22)	1		
Variety	−.33 (n = 15)	.84** (n = 16)	−.11 (n = 18)	1	
Quality	.57* (n = 15)	.007 (n = 13)	−.22 (n = 14)	−.38 (n = 12)	1

*$p < .05$, **$p < .01$.

generated. The presence of examples had a positive effect on *novelty* ($M = .28$, $SE = .10$) in that the exemplar group generated a larger number of novel ideas than the control group. Although the effect size on *quality* was not significantly different from zero, its confidence interval ranged between medium positive (.39) to small negative (−.23) suggesting that the presence of examples had moderate positive effect on the quality of solution ideas in some studies but negative effect in others. We will present the results of the analyses examining how the timing, the common-ness, and the number of examples modify the magnitude of the effects on examples on these five measures in the next section.

We also examined the correlation between the unbiased effect size estimates on these five measures (see Table 3). The *quantity* and *variety* of solution ideas were positively correlated with each other ($r = .84$, $p < .01$), suggesting that individuals producing more categories of ideas also generated more ideas. More critically, the *degree of copying* was negatively correlated with the *quantity* of solution ideas ($r = -.51$, $p = .008$) but positively correlated with the *quality* of ideas ($r = .57$, $p = .03$). A broad search in the problem space promotes the variety of ideas, while a narrow but deep search in the problem space reduces the variety but improves the quality of ideas (Rietzschel et al., 2007). Our correlation results suggest that the more individuals copy from the examples the narrower and deeper their solution search will be, and this will reduce the quantity but improve the quality of ideas.

In sum, the results converge to support that individuals' design problem solving processes will be influenced by the presence of the examples. Individuals will narrow down their scope of search focusing on example-related domains, and this will reduce the variety but enhance the quality and novelty of the solution ideas. The next section presents the results of the analyses examining how the timing, the common-ness, and the number of examples modify the magnitude of these exemplar effects.

3.2 The impact of timing, common-ness, and number of examples on effect sizes

For the 43 studies included in this meta-analysis, 27 of them presented the examples before problem solving ("before" studies), and the rest of them

Table 4 The random variance component, weighted mean, standard error, and 95% confidence interval of the unbiased effect size estimates in the "before" and "during" studies

	Degree of copying		*Quantity*		*Novelty*		*Variety*		*Quality*	
	Before	*During*	*Before*	*During*	*Before*	*During*	*Before*	*During*	*Before*	*During*
Number of studies	19	9	19	12	13	16	9	11	9	10
Common-ness Score M (SD)	3.76 (.79)	2.36 (.35)*	3.60 (.92)	2.30 (.35)*	3.78 (.97)	2.27 (.33)*	3.69 (1.15)	2.22 (.20)*	3.69 (1.15)	2.22 (.20)
Numbers of example M (SD)	1.19 (.60)	2.00 (.47)*	1.33 (.86)	2.46 (.97)*	1.73 (1.16)	2.35 (1.00)	1.78 (1.20)	1.91 (1.15)	1.78 (1.27)	1.91 (.53)
Random variance component	.02	.00	.00	.00	.16	.00	.00	.12	.27	.00
Weighted Effect size	.46	.22	.16	−.15	.24	.33	−.12	−.53	.08	.001
SE of the weighted effect size	.11	.20	.09	.18	.16	.13	.02	.22	.22	.23
95% confidence interval	.68, .24	.63, −.18	.33, −.02	.22, −.51	.56, −.08	.60, .05	−.08, −.15	−.50, −.95	.54, −.38	.48, −.48
"Before" vs. "During" (common-ness score and number of example treated as the covariates)	$F(1, 24) = 1.40$, $p = .25$		$F(1, 27) = 4.96$, $p = .03$		$F(1, 25) = .007$, $p = .94$,		$F(1,16) = 7.02$, $p = .02$		$F(1, 15) = 12.23$, $p = .003$	

* Significantly different from the "before" group, $p < .05$.

Table 5 Regression model for each measure

	Degree of copying		*Quantity*		*Novelty*		*Variety*		*Quality*	
	Before	*During*	*Before*	*During*	*Before*	*During*	*Before*	*During*	*Before*	*During*
Adjusted R^2	.26	<.001	.04	<.001	.26	.16	.71	.16	.81	.12
F-value	4.17	.29	1.36	.27	3.32	2.46	10.61	1.92	20.61	1.59
p-value	.04	.76	.29	.77	.08	.12	.01	.21	.001	.27
Predictors										
Number of examples Standardized coefficient β	−.53*	−.17	.30	.22	$-.52^{+}$	$-.47^{+}$	−.08	.56	.08	.25
Common-ness Standardized coefficient β	−.55*	−.19	−.18	.19	−.40	−.22	−.91*	.02	−.91*	.50

$^{*}p \leq .05$, $^{+}p \leq 0.08$.

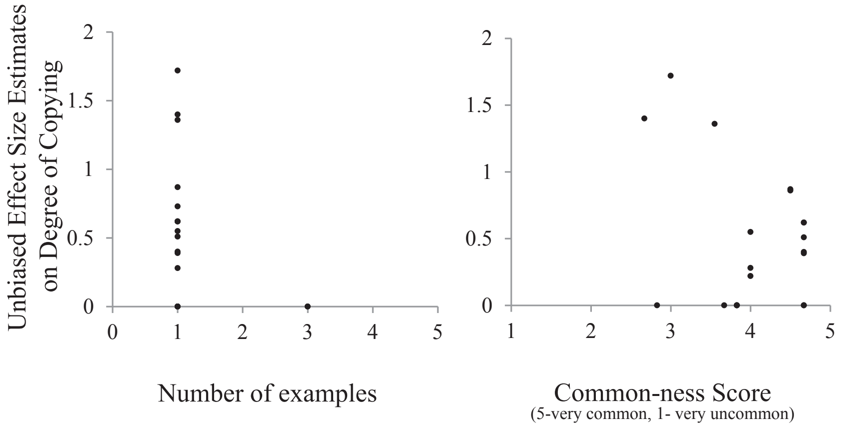

Figure 1 The scatter plot of the unbiased effect sizes against the number of examples (left panel) and the common-ness score (right panel) for the "before" studies

presented the examples during problem solving ("during" studies). The common-ness score of the examples presented in the "before" studies (common-ness score: $M = 3.82$, $SD = .88$) were higher than those presented in the "during" studies (common-ness score: $M = 2.22$, $SD = .22$), $t(41) = 7.12$, $p < .001$. This suggests that the examples presented in the "before" studies were more common. There was also a significant difference between these two subsets of studies on the number of examples presented ("before" studies: $M = 1.37$, $SD = .93$, "during" studies: $M = 2.44$, $SD = .96$), $t(41) = 3.59$, $p = .001$. Therefore, studies included in this meta-analysis were divided into two subsets: before vs. during, based on the timing of when the examples were presented. New weightings were computed for each of the subsets. Table 4 presents the random variance component, the weighted mean, standard error, and 95% confidence interval of the unbiased effect sizes in each subgroup. Weighted least-squares regression analyses

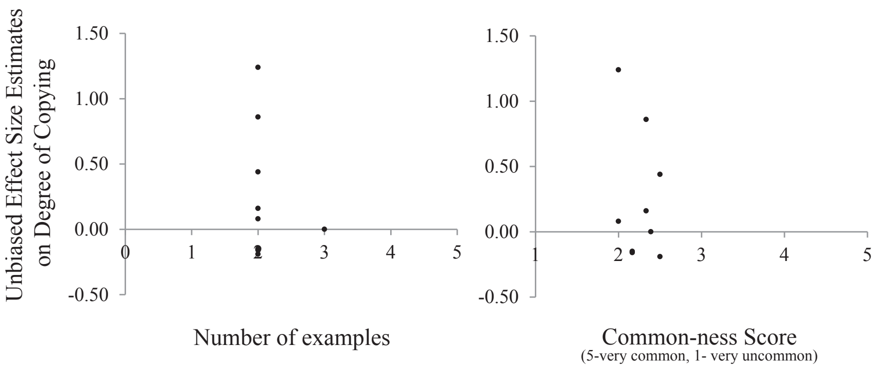

Figure 2 The scatter plot of the unbiased effect sizes against the number of examples (left panel) and the common-ness score (right panel) for the "during" studies

were carried out to identify the independent contribution of each of the two moderators: the number and the common-ness of the examples to degree of copying, quantity, novelty, variety, and quality of ideas reported in the "before" and "during" studies respectively (see Table 5 for the regression results).

3.2.1 Degree of copying

Exposure to examples at the beginning of and during problem solving both induced a moderate degree of copying ("before" studies: $M = .46$, $SD = .11$, "during" studies: $M = .22$, $SD = .20$, "before" vs. "during" studies: $F(1, 26) = 1.14$, $p = .30$).

For the "before" studies, the negative coefficient associated with *common-ness score* ($\beta = -.55$) in the regression model implies that individuals would copy less from the common examples. Figure 1 presents the scatter plot of the effect size against the common-ness score (right panel). The regression model also revealed a negative correlation between the number of examples and the degree of copying ($\beta = -.53$). However, most of the "before" studies have presented only one example to individuals and only two "before" studies have presented multiple examples and measured the degree of copying (see Figure 1, left panel, where both presented three examples and reported zero effect size). Due to the highly skewed distribution, the finding on the impact of *number of examples* is questionable.

The regression model for the "during" studies was not statistically significant, suggesting that the degree of copy was not impacted by the *common-ness score* and *number of examples*. The non-significant findings are likely due to the small variance on these two variables (see Figure 2).

In sum, the regression analyses revealed a moderate degree of copying induced by the examples ("before" studies: $M = .46$, $SD = .11$, "during" studies: $M = .22$, $SD = .20$). If individuals received the examples at the beginning of problem solving, they tended to copy more from uncommon than common examples. For studies presenting the examples after a period of initial work, the degree of copying was not affected by the number nor the common-ness of the examples. The absence of any moderating effect may be due to the small variance on these two variables in the "during" studies.

3.2.2 Quantity

The effect size on quantity for the "before" studies ($M = .16$, $SE = .09$) was larger than the "during" studies ($M = -.15$, $SE = .18$), $p = .03$. However, neither the effect size on quantity for the "before" nor the "during" studies was significantly different from zero. Also, the regression models were not statistically significant. Together, the very small effect sizes and the non-significant regression models converge to suggest that the presence of examples did not have any direct and significant impact on the quantity of ideas generated.

3.2.3 Novelty

Both presenting the examples at the beginning of and during problem solving generated a small-medium positive effect on novelty ("before" studies: $M = .32$, $SD = .44$; "during" studies: $M = .14$, $SD = .13$, "before" vs. "during" studies: $F(1, 32) = .002$, $p = .96$).

The regression models for the "before" and "during" studies were both marginally significant ("before" studies: $p = .08$, "during" studies: $p = .12$). Both regression models reported a negative coefficient on *number of examples* ("before" studies: $\beta = -.52$, $p = .05$; "during" studies: $\beta = -.47$, $p = .07$) in that presenting more examples resulted in smaller positive effect on novelty as compared to control. When examining the impact of the common-ness of examples on the novelty of the solution ideas, the regression model for the "before" studies revealed that participants benefitted more from uncommon examples than common ones, as indicated by the negative coefficient on *common-ness score* ($\beta = -.36$, $p = .16$). A similar trend was observed for the regression model for "during studies" ($\beta = -.22$, $p = .36$), however, the impact did not approach significance.

The results of regression models for the "before" and "during" studies were similar to each other. However, both regression models were just marginally significant, and we suggest that these may be due to the small number of studies included in the regression models. We therefore merged the "before" and "during" studies together and conducted another weighted least-squares regression analysis using *common-ness score*, *number of examples,* and *timing* as the predictor variables and the unbiased effect size estimate weighted by the inverse of its variance as the dependent variable. The regression model

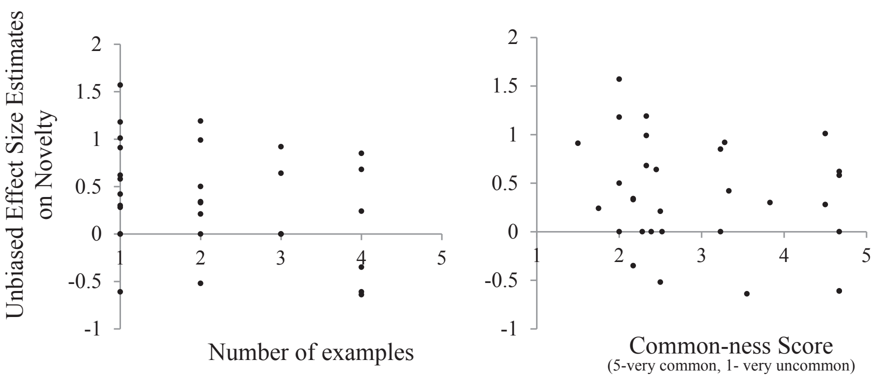

Figure 3 The scatter plot of the unbiased effect sizes against the number of examples (left panel) and the common-ness score (right panel)

was statistically significant, $F(3, 28) = 3.99$, $p = .02$, adjusted $R^2 = .24$. The coefficient on number of examples was negative (number of examples: $\beta = -.57$, $p = .005$) implying that the positive exemplar effect on novelty diminished as the number of examples increased (see Figure 3). The coefficient on *common-ness score* was also negative, $\beta = -.46$, but it was only marginally significant, $p = .10$. This suggests that increasing the common-ness of the examples also diminished the positive effect on novelty, but it was less impactful than increasing the number of examples presented.

Adding the interaction term between *number of examples* and common-ness *score* did not improve the predictive power of the model (adjusted R^2 dropped from .24 to .21), and the coefficient associated with the interaction term was not statistically significant ($\beta = -.02$, $p = .97$). These suggest the absence of any interaction effect between these two moderators. Also, the variable "timing" did not have significant impact on the effect size on variety ($\beta = -.02$, $p = .94$).

3.2.4 Variety

The effect sizes on the variety of solution ideas for "before" and "during" studies were both significantly smaller than zero, indicating that individuals in the exemplar condition generated fewer categories of ideas than those in the control condition. The negative impact was more significant when examples were presented during than at the beginning of problem solving ("before" studies: $M = -.12$, $SE = .02$; "during" studies: $M = -.53$, $SE = .22$, $p = .02$). One possible explanation is that after a period of initial work, individuals may have expanded their problem space, and the presented example solutions were more likely to be overlapped with their current problem solving approach, making individuals more fixated on the current approach.

The only significant coefficient in the regression model for "before" studies was *common-ness score* ($\beta = -.91$). The negative coefficient associated with *common-ness score* indicated that presenting less common examples could reduce

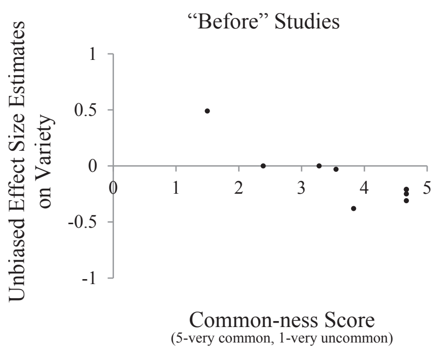

Figure 4 The scatter plot of the unbiased effect sizes on variety against the common-ness score for "before" studies

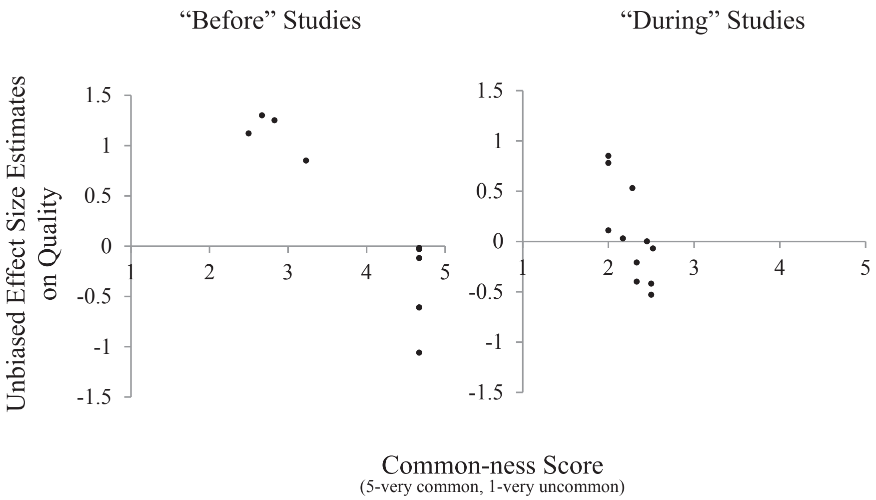

Figure 5 The scatter plot of the unbiased effect sizes against the common-ness score for "before" (left panel) and "after" studies (right panel)

the negative impact on the variety of ideas (see Figure 4). For "during" studies, neither the number nor the common-ness of the examples affected the effect size on variety.

3.2.5 Quality

For both "before" and "during" studies, the effect sizes on quality were neither all positive nor all negative, with large confidence intervals overlapped with zero ("before" studies: 54 to −.38, "during" studies: .48 to −.48). The weighted mean of effect sizes on quality for "before" studies ($M = .08$, $SE = .22$) was significantly larger than "during" studies ($M = .001$, $SE = .30$), $p = .003$, suggesting that, in general, individuals benefited more from examples presented at the beginning of than during problem solving. The large confidence intervals suggest the existence of moderators, affecting the effects of examples on the quality of solution ideas.

The weighted least-squares regression analysis on "before" studies reported a significant and negative impact of *common-ness score* on quality ($\beta = -.91$, $p < .001$) that increasing the common-ness of the presented examples can reverse the effect size on quality from positive to negative (see Figure 5, left panel). Yet, similar significant impact was not found for the "during" studies. However, there was only a small variance on the common-ness score among those "during" studies (see Figure 5, right panel), making it difficult to detect the impact of common-ness of the examples on solution quality.

The variable *number of examples* had no significant impact on the quality of solution ideas ("before" studies: $\beta = -.19$, $p = .25$; "during" studies: $\beta = -.59$, $p = .75$). However, the absence of any significant impact is likely due to the small variance on this variable. The majority of the "before" studies only presented one example to the individuals (8 out of 10), while most of the "during" studies presented two examples to the individuals (8 out of 10).

4 Discussion

This meta-analysis examined the effects of positive examples on design processes. The results indicate that individuals tend to copy ideas or parts of ideas from the examples and produced less variety of ideas, as compared to those who did not receive any example solution while solving the design task. The more individuals copy from the examples, the smaller number of ideas they can generate. However, looking at examples can significantly improve the quality and novelty of the solution ideas produced. These findings are consistent with our prediction that the presence of examples can modify the search strategy from a broad one to a focused one. Although this narrows the scope of search, it allows a more in-depth exploration, and in turn, improves solution quality.

When solving a design task, individuals may search randomly and broadly for different ideas at the beginning, and example solutions presented at the beginning of problem solving may serve as anchors and lead individuals to search in the example-related domains. In other words, the presence of examples should direct individuals to allocate their attention to example-related domains, facilitating a deeper search in those portions of the problem space. This in turn increases the likelihood of generating innovative and high-quality ideas. This explains the negative effects of examples on the variety of ideas and the positive effects on the novelty and quality of ideas.

According to this attention allocation explanation, individuals should benefit more from looking at one example than multiple non-negative examples. It is because multiple examples may activate different sets of concepts and encourage a multi-domain search, creating a more diffused rather than a focused search strategy. This may attenuate the role of examples in attention allocation. Consistent with this, our meta-analysis revealed that the number of examples was negatively correlated with the magnitude of the positive effect on the novelty of ideas.

The attention allocation explanation is also supported by our findings on the link between the common-ness of examples and the quality of solution ideas. If the role of examples is to help individuals select and confine a search space to explore, then looking at uncommon examples should lead individuals to search in less typical domains and in turn facilitate novel conceptual combination. Supporting this prediction, we found that individuals in the exemplar group generated higher novelty and quality solutions than those in the control group when the presented examples were uncommon.

Our analyses also reported a significant "timing" effect; presenting examples at the beginning of problem solving produced a larger positive impact on design solutions, as compared to presenting examples during problem solving. These results seem to be consistent with the prediction based on the sunk cost effect that after a period of initial work on the task, individuals have developed a commitment to their own approach; and this should lower their tendency to make use of the external information to improve the current problem solving approach.

Our results, somewhat surprisingly, do not offer a strong support to the open goal effect that goals which have been set but not completed have been shown to affect processing of external information in problem solving (Christensen & Schunn, 2007; Moss et al., 2007, 2011; Shah & Kruglanski, 2002). One possible mechanism underlying the open goal effect is that individuals may have expanded the problem space after a period of initial work, and it is, therefore, more likely for them to see the connection between the cues and the problem, and assimilate them to enhance problem solving performance, especially when the cues are novel and remote.

However, our meta-analysis did not report any significant impact of common-ness of examples on novelty and quality. We suggest that the absence of any open goal effect may be due to the small sample size as well as the small variance on the common-ness of examples presented during problem solving, making it difficult to detect the link between the common-ness of examples and novelty of the design solutions. Of note, however, Tseng et al. (2008) did demonstrate an open goal effect in that distant examples became more meaningful after work on the problem, where the expanded problem space likely enabled more distant examples to become meaningful. Said different, the "sweet spot" described by Fu and Chan et al. (2013) may shift as the problem space expands, in that case making for more impactful use of apparently distant examples.

Further, the sunk cost and open goal effects may not be mutually exclusive. Working on a problem should help expand the problem space, and this should increase the likelihood of assimilating novel and remote cues to enhance design solution; however, this should also increase the sunk cost (i.e. time) of design solution at the same time, making individuals more reluctant to modify the design. The sunk cost and open goal effects may counteract each other. This may nullify the benefits of examples. In order to leverage the effects of examples, especially for those encountered during design problem solving, we have to find ways to minimize the sunk cost effect. One possible way is to assign individuals to work on the design task briefly for multiple times rather than working on the task intensively for one single time. Although the actual cost (i.e., time) would be the same, approaching a problem in a distributed effort may lower the perception of cost, making individuals more likely to modify their problem solving strategy. Sio, Kotovsky, and Cagan (unpublished) reported that individuals solved more insight problems when the tasks were presented in a distributed rather than a sequential manner, especially for insight problem requiring a shift in search strategy.

In this meta-analysis, we found that individuals benefited more from looking at one single example than multiple ones. As discussed above, looking at multiple examples may activate different domains of concepts, each pointing to different search directions. It will be difficult to conduct a multi-domain search because of the limited cognitive resources humans possess. However, presenting multiple examples may be more beneficial in team design problem solving because teams consisting of multiple members have greater cognitive capacity and should be more likely to conduct a parallel and multi-domain search (see Knippenberg & Schippers, 2007 for a review). Until now, only a few studies examining the effects of examples on team design problem solving (Fu, Cagan, & Kotovsky, 2010) have been reported. Further studies can focus on comparing how presenting single vs. multiple examples impacts individual and team design processes.

In our analysis, we only focused on studies using example solutions that fit the design task requirements. Early studies of the effect of examples on design processes usually presented poor examples to individuals and examined if individuals were able to ignore the poor ideas embedded in the examples and generate good design solutions. These studies converged to report a strong negative effect of examples on design solution in that looking at examples could induce a large degree of copying and reduce the number of ideas generated. Will presenting good examples and bad examples generate a comparable degree of copying, or will individuals copy more from a good example because of its high relevancy to the problem? We compared the unbiased effect sizes on degree of copying between studies presenting non-negative and negative examples at the beginning of problem solving,[5] while controlling for the common-ness and number of examples. No significant difference on the degree of copying was found between these two sets of studies, non-negative examples: $M = .40$, $SD = .52$, negative examples: $M = .81$, $SD = .49$, $F(1, 38) = 1.33$, $p = .26$. This implies that copying from examples is a strong and stable phenomenon that will occur regardless the good/bad quality of the examples presented. However, it is important to note that copying does not necessarily lead to poor design solution ideas. A larger degree of copying may just imply that the examples induce a larger impact on individuals' problem solving approach, and it can lead individuals to search narrower and

5. Majority of the studies using negative examples presented the examples at the beginning of the study. There is no enough data for comparing the impact of negative and non-negative examples presented during problem solving.

deeper in the problem space, which enhances the quality and novelty of the design solutions.[6]

The majority of the participants of the studies included in this meta-analysis were engineering or industrial design seniors, who are considered to be at the early stage of expertise in solving design tasks. One may challenge that the effects of examples may be different on experienced designers because they are experienced in solving various design tasks so that they do not need to heavily rely on examples to frame the design task. Therefore, they should be less likely to be fixated and affected by examples. However, if design fixation is the result of the automatic activation of concepts by examples, then experts who have a well-structured and easily-activated domain-specific knowledge should find it more difficult to avoid fixation. There are studies showing that both experts and students in mechanical engineering demonstrated a significant degree of design fixation when solving design problems (Jansson & Smith, 1991; Linsey et al., 2010; Viswanathan & Linsey, 2013a). Therefore, we predict that the presence of examples should impact the experts' and novices' design processes in a similar way that it can reduce variety but enhance the quality and novelty of the solution ideas. However, we predict that the effect of the moderators – the common-ness and number of examples – may have differential impact on novices and experts. First, the common-ness of an example depends on the amount of domain-specific knowledge an individual possesses. A distant example may be beneficial for experts who have a large amount of domain-specific knowledge (Chi, Feltovich, & Glaser, 1981). However, presenting the same example to novices may not yield positive effects because novices may not have the knowledge to see the connection between the example and the problem. Moss, Kotovsky, and Cagan (2006) revealed that engineering freshmen were less capable in seeing the connection between different concepts, as compared to engineering seniors. The low connectivity between the concepts may inhibit novices to make use of the examples effectively. Second, we predict that experts may be more likely than novices to benefit from multiple examples due to the expert-novice difference on problem solving approach. Experts process information at a higher level of abstraction (Chi et al., 1981), and this allows them to abstract the examples and process the large amount of information embedded in the multiple examples efficiently. They may also have more automated move operators for incrementing designs that impose less of a cognitive load (Kotovsky, Hayes & Simon, 1985). Also, when solving a design task, experts tend to take a breadth-first approach where they gather a large and wide range of related-information at the beginning of problem solving while novices tend to do a depth-first search (Cross, 2004).

6. It is impossible to examine if the presence of negative examples can significantly impact the quality and novelty of the design solutions because the majority of studies using negative examples do not measure novelty and quality of the design solutions.

Looking at multiple examples facilitates a broad search in the problem space, and this is consistent with experts' general problem solving approach. Experts should therefore be more likely than novices to benefit from the presence of multiple examples.

It should be noted that, despite efforts to include studies from different sources, the number of studies included in this meta-analysis was of necessity, relatively small for typical regression analysis. It is likely that our analysis is underpowered to detect all the exemplar effects. A few of the regression models reported in this paper are just marginally significant; with more studies, these models could reach statistical significance. However, even though we only have a small number of studies included in the meta-analysis, our findings converge to support that the presence of examples can modify the search strategy from a diffused search to a focused one, and this can significantly impact the variety, novelty, and quality of the design solutions.

Our meta-analysis also identifies some limitations across the previous studies. Previous studies examining early exposure to examples usually presented one single common example while those examining late exposure to examples presented multiple uncommon examples (see Table 3). The small overlap between the characteristics of examples used in these two sets of studies makes it difficult to fully understand how the common-ness and the number of examples impact individuals' search strategy at different stages of design problem solving. To address this, studies with a large variance in terms of the common-ness and number of examples should be conducted.

In our study, we did not report a strong correlation between the quantity and quality of the solution ideas ($r = .007$, $p > .20$). This challenges the common assumption among the brainstorming studies that the greater the number of ideas generated, the greater the chance of producing a good solution (Osborn, 1963). According to our meta-analysis, generating high quantity and generating high quality of ideas required different search strategies. Looking at examples led individuals to explore the problem space more narrowly and deeply, and this promoted the quality and novelty of the design solution. These support the use of examples in the design process because, in real-life problem solving, people usually value quality over quantity of the solutions (Rowatt, Nesselroade, Beggan, & Allison, 1997). Providing a single rather than multiple examples allows individuals to focus on fewer domains and to search deeper. An uncommon example will direct individuals to search in a less typical domain. Together, a focused search in an uncommon domain should facilitate novel conceptual combination. These support our prediction that looking at examples can impact how individuals explore the problem space, and this can potentially facilitate the design processes, depending on the number, common-ness, and the timing of the presentation.

Acknowledgments

This research was supported by AFOSR Grant FA9550-12-1-0374.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.destud.2015.04.004

References

Agogué, M., Kazakçi, A., Hatchuel, A., Masson, P., Weil, B., Poirel, N., et al. (2014). The impact of type of examples on originality: explaining fixation and stimulation effects. *The Journal of Creative Behavior, 48*(1), 1–12. http://dx.doi.org/10.1002/jocb.37.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*(1), 124–140. http://dx.doi.org/10.1016/0749-5978(85)90049-4.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and non-anxious individuals: a meta-analytic study. *Psychological Bulletin, 133*(1), 1. http://dx.doi.org/10.1037/0033-2909.133.1.1.

Cagan, J., Kotovsky, K., & Simon, H. A. (2001). Scientific discovery and inventive engineering design: cognitive and computational similarities. In *Formal engineering design synthesis* (pp. 442–465). Cambridge University Press. http://dx.doi.org/10.1017/cbo9780511529627.016.

Cardoso, C., & Badke-Schaub, P. (2011). The influence of different pictorial representations during idea generation. *The Journal of Creative Behavior, 45*(2), 130–146. http://dx.doi.org/10.1002/j.2162-6057.2011.tb01092.x.

Cardoso, C., Gonçalves, M., & Badke-Schaub, P. (2012). Searching for inspiration during idea generation: pictures or words? in *DS 70*. In *Proceedings of design 2012, the 12th international design conference, Dubrovnik, Croatia*.

Casakin, H., & Goldschmidt, G. (1999). Expertise and the use of visual analogy: Implications for design education. *Design Studies, 20*(2), 153–175. http://dx.doi.org/10.1016/s0142-694x(98)00032-5.

Chan, J., Dow, S. P., & Schunn, C. D. (2014). Do the best design ideas (really) come from conceptually distant sources of inspiration? *Design Studies, 36*, 31–58. http://dx.doi.org/10.1016/j.destud.2014.08.001.

Chan, J., Fu, K., Schunn, C., Cagan, J., Wood, K., & Kotovsky, K. (2011). On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical distance, commonness, and modality of examples. *Journal of Mechanical Design, 133*(8), 081004. http://dx.doi.org/10.1115/1.4004396.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121–152. http://dx.doi.org/10.1207/s15516709cog0502_2.

Chiu, I., & Shu, H. (2012). Investigating effects of oppositely related semantic stimuli on design concept creativity. *Journal of Engineering Design, 23*(4), 271–296. http://dx.doi.org/10.1080/09544828.2011.603298.

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. http://dx.doi.org/10.3758/bf03195939.

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. http://dx.doi.org/10.1037/0278-7393.31.5.1134.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.

Cross, N. (2004). Expertise in design: an overview. *Design Studies, 25*(5), 427–441. http://dx.doi.org/10.1016/j.destud.2004.06.002.

Dahl, D. W., & Moreau, P. (2002). The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research, 39*(1), 47–60. http://dx.doi.org/10.1509/jmkr.39.1.47.18930.

Dodds, R. A., Smith, S. M., & Ward, T. B. (2002). The use of environmental clues during incubation. *Creativity Research Journal, 14*(3–4), 287–304. http://dx.doi.org/10.1207/s15326934crj1434_1.

Duncker, K. (1945). On problem-solving. *Psychological Monographs, 58*(5), i-113. http://dx.doi.org/10.1037/h0093599.

Fu, K., Cagan, J., & Kotovsky, K. (2010). Design team convergence: the influence of example solution quality. *Journal of Mechanical Design, 132*(11), 111005. http://dx.doi.org/10.1115/1.4002202.

Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C., & Wood, K. (2013). The meaning of "near" and "far": the impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design, 135*(2), 021007. http://dx.doi.org/10.1115/detc2012-70420.

Fu, K., Murphy, J., Yang, M., Otto, K., Jensen, D., & Wood, K. (2013). Investigating the effect of functionality level of analogical stimulation on design outcomes. In *Design engineering workshop, Kitakyushu, Fukuoka, Japan*.

Goel, A. K. (1997). Design, analogy, and creativity. *IEEE Expert, 12*(3), 62–70. http://dx.doi.org/10.1109/64.590078.

Goel, V. (1995). *Sketches of Thought*. Cambridge, MA: MIT Press.

Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2012). *Find your inspiration: exploring different levels of abstraction in textual stimuli*. Paper presented at the 2nd international conference on design creativity (ICDC2012), Glasgow UK.

Hassard, S. T., Blandford, A., & Cox, A. L. (2009). Analogies in design decision-making. In *Proceedings of the 23rd British HCI group annual conference on people and computers: celebrating people and technology* (pp. 140–148). British Computer Society.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Helms, M., Vattam, S. S., & Goel, A. K. (2009). Biologically inspired design: process and products. *Design Studies, 30*(5), 606–622. http://dx.doi.org/10.1016/j.destud.2009.04.003.

Hey, J., Linsey, J., Agogino, A. M., & Wood, K. L. (2008). Analogies and metaphors in creative design. *International Journal of Engineering Education, 24*(2), 283.

Holyoak, K. J., & Thagard, P. (1995). *Mental Leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies, 12*(1), 3–11. http://dx.doi.org/10.1016/0142-694x(91)90003-f.

Kazakci, A. O., Gillier, T., Piat, G., & Hatchuel, A. (2014). Brainstorming versus creative design reasoning: a theory-driven experimental investigation of novelty, feasibility and value of ideas. *Design Computing and Cognition* 1−20.

Kontopantelis, E., Springate, D. A., & Reeves, D. (2013). A re-analysis of the Cochrane Library data: the dangers of unobserved heterogeneity in meta-analyses. *PloS one, 8*(7), e69930. http://dx.doi.org/10.1371/journal.pone.0069930.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? evidence from Tower of Hanoi. *Cognitive Psychology, 17*(2), 248−294. http://dx.doi.org/10.1016/0010-0285(85)90009-x.

Linsey, J. S., Tseng, I., Fu, K., Cagan, J., Wood, K. L., & Schunn, C. (2010). A study of design fixation, its mitigation and perception in engineering design faculty. *Journal of Mechanical Design, 132*(4), 041003. http://dx.doi.org/10.1115/1.4001110.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lujun, Z. (2011). Design fixation and solution quality under exposure to example solution. In *2011 IEEE 2nd International Conference on computing, control and industrial engineering (CCIE)*. http://dx.doi.org/10.1109/ccieng.2011.6007974.

Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine, 20*, 641−654. http://dx.doi.org/10.1002/sim.698.

McComb, C., Cagan, J., & Kotovsky, K. (2015). Rolling with the punches: an examination of team performance in a design task subject to drastic changes. *Design Studies, 36*, 99−121. http://dx.doi.org/10.1016/j.destud.2014.10.001.

Moss, J., Kotovsky, K., & Cagan, J. (2006). The role of functionality in the mental representations of engineering students: some differences in the early stages of expertise. *Cognitive Science, 30*(1), 65−93. http://dx.doi.org/10.1207/s15516709cog0000_45.

Moss, J., Kotovsky, K., & Cagan, J. (2007). The influence of open goals on the acquisition of problem-relevant information. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*(5), 876−891. http://dx.doi.org/10.1037/0278-7393.33.5.876.

Moss, J., Kotovsky, K., & Cagan, J. (2011). The effect of incidental hints when problems are suspended before, during, or after an impasse. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 140. http://dx.doi.org/10.1037/a0021206.

Osborn, A. F. (1963). *Applied Imagination: Principles and procedures of creative problem-solving* (3rd ed.). New York: Scribner's.

Perttula, M., & Sipilä, P. (2007). The idea exposure paradigm in design idea generation. *Journal of Engineering Design, 18*(1), 93−102. http://dx.doi.org/10.1080/09544820600679679.

Purcell, A. T., & Gero, J. S. (1992). Effects of examples on the results of a design activity. *Knowledge-Based Systems, 5*(1), 82−91. http://dx.doi.org/10.1016/0950-7051(92)90026-c.

Purcell, A. T., & Gero, J. S. (1996). Design and other types of fixation. *Design Studies, 17*(4), 363−383. http://dx.doi.org/10.1016/s0142-694x(96)00023-3.

Purcell, A. T., Williams, P., Gero, J. S., & Colbron, B. (1993). Fixation effects: do they exist in design problem solving? *Environment and Planning B, 20*, 333−345. http://dx.doi.org/10.1068/b200333.

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2007). Relative accessibility of domain knowledge and creativity: the effects of knowledge activation on the quantity and originality of generated ideas. *Journal of Experimental Social Psychology, 43*(6), 933−946. http://dx.doi.org/10.1037/e633982013-064.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192. http://dx.doi.org/10.1037/0033-2909.118.2.183.
Rowatt, W. C., Nesselroade, K., Beggan, J. K., & Allison, S. T. (1997). Perceptions of brainstorming in groups: the quality over quantity hypothesis. *The Journal of Creative Behavior, 31*(2), 131–150. http://dx.doi.org/10.1002/j.2162-6057.1997.tb00786.x.
Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: how accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology, 38*(4), 368–383. http://dx.doi.org/10.1016/s0022-1031(02)00005-7.
Shah, J. J., Smith, S. M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. *Design Studies, 24*(2), 111–134. http://dx.doi.org/10.1016/s0142-694x(02)00034-0.
Simon, H. A. (1969). *The sciences of the artificial, Vol. 136*. MIT press.
Sio, U. N., Kotovsky, K., & Cagan, J. (2015). *Working smarter but not longer: Achieving insight without impasse*. (unpublished).
Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 135*(1), 94–120. http://dx.doi.org/10.1037/a0014212.
Sio, U. N., & Ormerod, T. C. (2014). Incubation and cueing effects in problem-solving: set aside the difficult problems but focus on the easy ones. *Thinking & Reasoning, 21*(1), 113–129. http://dx.doi.org/10.1080/13546783.2014.886626.
Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition, 21*(6), 837–845. http://dx.doi.org/10.3758/bf03202751.
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. http://dx.doi.org/10.1016/j.destud.2008.01.003.
Viswanathan, V. K., & Linsey, J. S. (2013a). Design fixation and its mitigation: a study on the role of expertise. *Journal of Mechanical Design, 135*(5), 051008. http://dx.doi.org/10.1080/09544828.2014.885934.
Viswanathan, V. K., & Linsey, J. S. (2013b). Role of sunk cost in engineering idea generation: an experimental investigation. *Journal of Mechanical Design, 135*(12), 121002. http://dx.doi.org/10.1115/1.4025290.
Ward, T. B. (2007). Creative cognition as a window on creativity. *Methods, 42*(1), 28–37. http://dx.doi.org/10.1016/j.ymeth.2006.12.002.
Williams, W. M., & Sternberg, R. J. (1988). Group intelligence: why some groups are better than others. *Intelligence, 12*(4), 351–377. http://dx.doi.org/10.1016/0160-2896(88)90002-5.
Wilson, J. O., Rosen, D., Nelson, B. A., & Yen, J. (2010). The effects of biological examples in idea generation. *Design Studies, 31*(2), 169–186. http://dx.doi.org/10.1016/j.destud.2009.10.003.