Overreliance on conceptually far sources decreases the creativity of ideas

Joel Chan (joc59@pitt.edu), Christian Schunn (schunn@pitt.edu)
Learning Research and Development Center, University of Pittsburgh
3939 O'Hara St, Pittsburgh, PA 15260, USA

Steven Dow (spdow@cs.cmu.edu)
Human Computer-Interaction Institute, Carnegie-Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213, USA

Abstract

Ideas are often generated from inspiration sources (e.g., prior experiences with the world, solutions to analogous problems). These sources may have benefits but also pitfalls (e.g., difficulty thinking of alternative approaches). In this paper, we investigate whether and how features of inspiration sources predict their impact on creative outcomes. In particular, we examine the popular but unevenly supported hypothesis that conceptually distant sources of inspiration provide the best insights for creative production. We test this hypothesis in the context of a Web-based real-world creativity platform, while addressing key methodological issues in prior empirical studies (e.g., truncated time scale, low statistical power, problem variation). Through a text analysis of many hundreds of concepts, we test whether greater conceptual distance between a concept's cited sources and the problem domain increases its probability of creative success (in this case, being shortlisted by an expert panel as a promising creative concept). We found that concepts that cite sources had greater success than those that did not cite sources of inspiration. However, increases in mean conceptual distance of sources actually decreased the probability of success, suggesting that far sources do not uniquely boost creativity and that an overreliance on far sources may even harm creativity. This negative effect of distance was robust across authors and different design problems on the platform. In light of these findings, we revisit theories of creative inspiration and general creative cognition.

Keywords: Creativity; analogy; problem-solving; in-vivo

Introduction

In the creative process, people inevitably build new ideas from sources of inspiration, most often from their prior knowledge and experiences (Ward, 1994). These sources of inspiration can lead one astray — e.g., incorporating undesirable features from existing solutions (Jansson & Smith, 1991), difficulty thinking of alternative approaches (Wiley, 1998) — but sometimes drive creative breakthroughs (Eckert & Stacey, 1998; Hargadon & Sutton, 1997). Are there features of inspirations that can predict when they will harm or hurt creativity? One potential feature of interest is the conceptual distance of those sources from one's working domain. For instance, consider the problem of e-waste accumulation: the world generates 20-50 million metric tons of e-waste every year, yielding environmentally hazardous additions to landfills. An innovator might approach this problem by building on near sources like smaller-scale electronics reuse/recycle efforts, or by drawing inspiration from a far source like edible food packaging technology (e.g., to design re-usable electronics parts). The central question we consider here is: what are the relative benefits of different levels of source conceptual distance for creative outcomes?

Many authors, principally those studying the role of analogy in the creative process, have proposed that conceptually far sources — structurally similar ideas with many surface (or object) dissimilarities (e.g., the atom/solar system analogy) — have more potential to yield more creative ideas (Gentner & Markman, 1997; Holyoak & Thagard, 1996; Ward, 1998). Empirically, the literature provides a mixed picture. A number of studies have shown an advantage of far over near sources for quality and flexibility of ideation, in addition to novelty of ideas (Chan et al., 2011; Chiu & Shu, 2012; Dahl & Moreau, 2002; Gonçalves et al., 2013; Hende et al., 2002). However, some *in vivo* studies of creative discovery have failed to find strong connections between far sources and creative mental leaps (Chan & Schunn, accepted), and other experiments have demonstrated equivalent benefits of both far and near sources for creative outcomes (Enkel & Gassmann, 2010; Malaga, 2000; Tseng et al., 2008), and even harmful effects of distance on creative outcomes (Fu et al., 2013)

What does this imply for theories of creative inspiration? Perhaps it is necessary to abandon or revise the theory that far sources uniquely support creativity, in line with theorists like Perkins (1983), who argues that conceptual distance does not matter, or Weisberg (2009), who argues that within-domain expertise is a primary driver of creativity. However, this assumes that opposing findings have a strong empirical foundation. Here, we argue that this is not the case: there are key methodological shortcomings in prior work that should be addressed before considering theory revision.

One potential methodological shortcoming is that the length of prior studies (typically 30 minutes to 1 hour problem-solving time) may be too short to observe the potential long-term payoffs of cross-domain inspiration. Scarce cognitive resources are required to ignore irrelevant surface details and attend to potentially insightful structural similarities. This might partially explain losses in fluency sometimes observed with the use of far sources (Chan et al., 2011; Hender et al., 2002). Problem solvers may be unwilling or unable to pay these higher relative costs of processing far sources in the context of a short task, whereas the processing cost would be reasonable on a more realistic design time scale. For example, 20 minutes is a substantial cost when the ideation phase is 1 hour long, but a negligible cost in a time span of weeks/months (a more realistic time scale for ideation phases in typical design projects). Relatedly, there may be low expected returns with few samples of low probability/high gain choices. At shorter time scales, creators might not have enough samples to consistently find these “hidden gems” for maximal inspirational payoff.

An additional issue in prior studies is a lack of statistical power. Among existing experimental studies, most have an N of 12 or less per treatment cell (Chiu & Shu, 2012; Hender et al., 2002; Malaga, 2000); only 4 studies had an N of 18 or better per cell (Chan et al., 2011; Fu et al., 2013; Gonçalves et al., 2013; Tseng et al., 2008), and they are evenly split in support/opposition for the benefits of far sources. Among the few correlational studies, only Dahl and Moreau (2002) had an acceptable study design in this regard, with 119 participants and a reasonable range of conceptual distance. Enkel and Gassmann (2010) only sampled 25 cases, and suffered from range restriction because they only sampled cases of cross-industry transfer.

Thus, the mixed empirical evidence base may reflect the proliferation of false negatives due to insufficient statistical power (potentially exacerbated by small or potentially zero effects at short time scales); on the other hand, the underpowered designs may have also yielded severe overestimation of effect sizes (i.e., false positives; Button et al., 2013).

A final methodological problem has to do with problem variation. Many of the experimental studies focus on a single design problem. It could be that some of the inconsistency of outcomes is the result of some design problems having unique characteristics.

The current work addresses all of these methodological issues (time scale, statistical power, problem variation) to yield stronger evidence to guide theorizing about the impact of conceptual distance on creative outcomes.

Methods

Research Context

The current work is conducted in the context of OpenIDEO (www.openideo.com), a large-scale Web-based crowdsourced open innovation platform that addresses various social problems (e.g., managing e-waste, increasing accessibility in elections). Problems are sponsored by an external company/organization, and instantiated as OpenIDEO *challenges*. Challenges begin with presentation of the *challenge brief*, crafted collaboratively with OpenIDEO designers, which gives a broad overview of the problem to be solved.

Over the subsequent ~10 weeks, contributors to the platform first post *inspirations* (e.g., descriptions of solutions to analogous problems, case studies of stakeholders) for a given problem, which help to define the problem space and identify promising solution approaches, and then *concepts*, i.e., specific solutions to the problem. Concepts are typically ~150 words long, providing more detail than one or two words/sentences/sketches, but less detail than a full-fledged design report (see Fig. 1 for an example concept). When posting concepts, contributors are prompted to cite inspirations that serve as sources of inspiration. These cited sources are stored and displayed as metadata for the concept.

> Lease electronics. Companies with electronics that fit a requirement of durability and sustainability should be made 'leasable'. And at the end of the devices' usability, it will then be returned to the manufacturing company. How many people have changed their current electronic models for a newer version after using it for approximately 1-2 years, even though the electronic device still works perfectly? Answer to this? Electronics should be leasable. But not all electronics. Only electronics that have been proven to have long-lasting durability and possesses the ability to be recycled can be eligible to be recycled. Once the electronic is leased, when it is returned, it can either be leased out again (dependent on condition) or it will be returned to the original manufacture to be recycled. Benefits to the consumers: We're constantly replace our devices anyways, and the cost of leasing would be lower than the cost of purchasing. Also, by purchasing a leased good, it ensures that there will be someone taking back the product at the end of the day, avoiding the responsibility of having to sell it. The consumer will also know that, since their product is leasable, it means that it can be recyclable. Benefits to the business leasing the electronics: The company now has an additional revenue stream in addition to selling devices. Also, by being able to lease the product out again and again, it increases the usage lifespan that a devices sees while still being able to keep inventory levels low for the company (don't need as many devices as they know that it will be returned after x-months). Benefits to the manufacturing company: They now have an incentive to create longer lasting goods as they know that consumers will be looking for leased products. Also, leasing ensures that the electronics will be returned to the manufacturer as opposed to ending up in the landfill. Providing further benefit of making sure the product is recyclable.

Figure 1: Raw text from example concept

Throughout each challenge, contributors give feedback on each other’s inspirations and concepts, primarily in the form of comments that are displayed on each inspiration/concept. A subset of concepts is shortlisted by an expert panel (composed of challenge sponsors, who are domain experts, and expert designers from OpenIDEO) for further refinement, taking both the novelty and feasibility/quality of each concept into consideration. A subset of the refined shortlisted concepts is then selected for real-world implementation.

Data Collection and Sample

The full dataset consists of 2,341 concepts posted for 12 completed challenges by 1,190 unique contributors (majority designers, domain experts), citing 4,557 unique inspirations; 110 of the concepts are shortlisted. These concepts and inspirations exist as public webpages on the OpenIDEO site, and were downloaded with OpenIDEO administrators’ permission.

Using a simple HTML parser, we extracted the full-text description of each concept/inspiration (for measurement of conceptual distance), and for all concepts, 1) information on what sources were cited, 2) number of comments received, and 3) an indicator for whether the concept was shortlisted for development.

Not all concepts cited inspirations as sources. Of the 2,341 concepts, 707 (posted by 357 authors) cited at least one inspiration, collectively citing 2,245 unique inspirations. 110 of these concepts (~16%) were shortlisted. This set of 707 concepts is the primary sample for this study; the others serve as a contrast to examine the value of explicit building at all on prior sources.

Measures

Conceptual Distance The unique nature of our dataset presented some methodological challenges to measuring distance. The complex and multifaceted nature of the various design problems made it difficult to distinguish between "within" and "between" domain sources in a consistent and principled manner. Continuous distance measures were an attractive alternative, but were too costly to obtain from human raters due to the large number of sources. Even with sufficient time, we were concerned about rater fatigue, possibly leading to poor reliability or drift in rating standards. To address these challenges, we employed a computational approach to measuring distance.

We used Latent Dirichlet Allocation (LDA; Blei et al. 2003) to learn a high-dimensional topic space from the full-text descriptions of the challenge briefs and concepts/inspirations. This approach is similar to Latent Semantic Analysis (Landauer et al., 1998), but for some purposes can produce stronger results (Griffiths et al., 2007). Briefly, LDA uses Bayesian probabilistic modeling to infer latent topics that produce the words in a given text based on statistical patterns of word use across texts in the corpus; similarity between texts is the degree of overlap in the texts' topics. To reduce potential noise, we first removed stopwords (e.g., "the", "which") from the texts. 400 topics were inferred from the entire collection of 6,910 documents. We then computed cosine similarity between each inspiration and its challenge brief when projected into this topic space.

For validation, five judges all coded continuous similarity (on a 1 to 6 scale) for 199 inspirations from one challenge. Although the task was difficult, the mean ratings across raters had an acceptable aggregate consistency intra-class correlation coefficient of .74. The LDA-based cosines correlated well with the human similarity ratings, $r = .51$, $p < .01$. This level of match was actually better than the highest pairwise agreement between the judges, reinforcing the value of automatic coding methods for this difficult task.

Fig. 2 shows examples of a near and far inspiration, along with the top 3 LDA topics (represented by the top 5 words associated with that latent topic), computed cosine vs. its challenge brief, and human similarity rating. Both inspirations are from the e-waste challenge, addressing the problem illustrated in the introduction above. For reference, the top 3 topics for the challenge brief are *{waste, e, recycling, electronics, electronic}*, *{waste, materials, recycling, recycled, material}*, and *{devices, electronics, electronic, device, products}* (distinguishing e-waste, general recycling, and electronics products)

The challenge briefs varied in length and specificity across challenges, as did mean raw cosines. To test if mean differences between challenges were meaningful, we compared cosines for 80 concepts from 4 challenges (whose mean raw cosines were very different from each other: 2 high and 2 low) with human judgments (coded separately but in the same way as above). The difference in mean cosine for the high vs low challenges was much smaller for the human judgments, $d = 0.18$, 95% CI = [–0.05 to 0.43], than the cosines, $d = 1.90$, 95% CI = [1.85 to 1.92], suggesting that between-challenge differences might be more an artifact of variance in challenge brief length/specificity. Thus, to ensure meaningful comparability across challenges, we normalized the cosines by computing the z-score for each inspiration's cosine relative to other inspirations from the same challenge. Similar results were obtained with raw cosines, but with more uncertainty in the estimates.

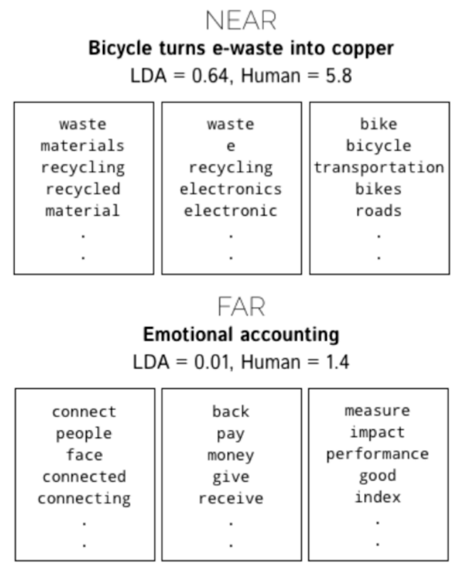

Figure 2: Topics found by LDA within examples of near and far inspirations for the e-waste challenge

To convert the cosines into a distance measure, we subtracted the cosine z-score from zero (so that higher = more distant). Then, each *concept's* conceptual distance measure was the mean distance of its cited inspirations.

Creative Outcomes Each concept's creative outcome measure was the binary status of whether or not it was shortlisted for further refinement. We note that this measure arises from the deliberations of an expert panel, a gold standard for measurement of creativity (Amabile, 1982), and combines consideration of both novelty and quality, the standard definition of creativity (Sawyer, 2009): the challenges are novel and unsolved, so by definition solving them would involve concepts that are different from (and, perhaps more importantly, significantly better than) existing unsatisfactory solutions.

Control measures Feedback (particularly raising questions/issues, suggestions for improvement) has the potential to significantly enhance the quality of the concept. Further, feedback may be an alternate pathway to success via source distance: building on far sources may attract more attention and therefore higher levels of feedback, thereby improving the concept; failing to account for feedback may lead to inflated estimates of the effects of source distance. Thus, we include the *number of comments* as a control measure.

Concepts could also cite other concepts as sources. Building on other highly creative concepts (measured by their shortlist status) could also significantly enhance the creativity of the concept. Thus, we also include the *number of cited shortlisted concepts* as an additional control measure.

Results

Descriptive Statistics

Because of our normalization procedure, the mean distance of cited inspiration sources was very close to 0, but ranged considerably. Feedback and number of shortlisted sources also varied considerably, particularly feedback.

Table 1: Descriptive statistics for predictor variables.

Variable	*M*	*SD*	Min	Max
Mean norm. distance	–0.10	0.85	–3.85	1.67
Feedback	8.43	9.45	0	67
Shortlisted sources	0.51	0.96	0	11

Statistical Models

Single-level Model We first fitted a logistic regression model with *shortlist* as the binary outcome, and *mean distance*, *feedback*, and *shortlisted sources* as predictors. The model estimates a *negative* effect of mean distance on shortlist probability – hereafter termed Pr(shortlist) — with a 1-unit increase in mean distance predicting a decrease of .38 in the log odds of being shortlisted (see Table 2). As an example, a concept with mean feedback and shortlisted sources, and mean distance = 0 would have predicted Pr(shortlist) = 0.13; increasing its mean distance to 1 would give predicted Pr(shortlist) = 0.09.

The model had a superior fit to a null model with no predictors, likelihood ratio = 73.50, $p < .001$ (for χ^2 with $df = 3$). Removing mean distance significantly increased model deviance by 7.78, $p < .01$ (for χ^2 with $df = 1$).

Table 2: Coefficient estimates for logistic regression of shortlist on mean distance, feedback, and shortlisted sources

	β	*SE*	exp(β)
Intercept	–2.66	0.18	0.07
Mean norm. distance	–0.38	0.14	0.68
Feedback	0.08	0.01	1.08
Shortlisted sources	0.16	0.10	1.17

Multilevel Models Given the multilevel structure of the data (concepts nested within authors, and also cross-classified within challenges), we also explored multilevel versions of the same model. Due to sample size restrictions (many missing cases and low n_{ij} for the crossed cells), we fitted 2 separate multilevel models: 1) a fixed-effects 2-level model with the same predictor specifications at level 1 (the concept level), and modeling author-level variation in mean Pr(shortlist), and 2) a random-effects 2-level model with the same predictor specifications at level-1, and modeling challenge-level variation in both mean Pr(shortlist), and the slope for the effect of mean distance.

The structure of these models is very similar, with the following structure at level 1:

$$\Pr(shortlist)_{ij} = \beta_{0j} + \beta_{1j}(DIST) + \beta_{2j}(FEEDBACK) \\ + \beta_{2j}(SHORTSOURCE)$$

where β_{0j} is the mean Pr(shortlist) for the j^{th} level-2 unit, $\beta_{1j}(DIST)$ is the estimated effect of mean normalized distance, $\beta_{2j}(FEEDBACK)$ is the estimated effect of feedback, and $\beta_{2j}(SHORTSOURCE)$ is the estimated effect of number of shortlisted sources.

The level-2 structure is:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

where γ_{00} is the grand mean Pr(shortlist) for all concepts; and u_{0j} is the level-2 variance in β_{0j}.

For modeling the challenge-level variation, we add the following structure at level-2:

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

where γ_{10} is the estimated effect of mean distance across challenges, and u_{1j} is the challenge-level variance in β_{1j}.

The author-multilevel model addresses potential concerns that estimates of the effect of mean distance may be untrustworthy due to failure to account for within-author similarity of concepts. The challenge-multilevel model addresses our key methodological interest in investigating robustness of effects across different problems. Table 3 summarizes the estimates and fits of these models in comparison to the simpler model discussed above.

Table 3: Coefficient estimates for the effect of mean normalized distance and model fit statistics with no nesting, author-nesting, and challenge-nesting

	Mean norm. distance	Deviance	AIC
No nesting	–0.37	537.74	545.74
Author-nesting	–0.40	536.14	546.14
Challenge-nesting	–0.36	507.02	517.02

Modeling author-level nesting improved fit by a small amount. Modeling challenge-level nesting considerably improved fit relative to the no-nesting model, with the lowered Akaike information criterion (AIC) suggesting that this was not due to overfitting. Importantly, the coefficients for mean distance remained substantially similar across all models,

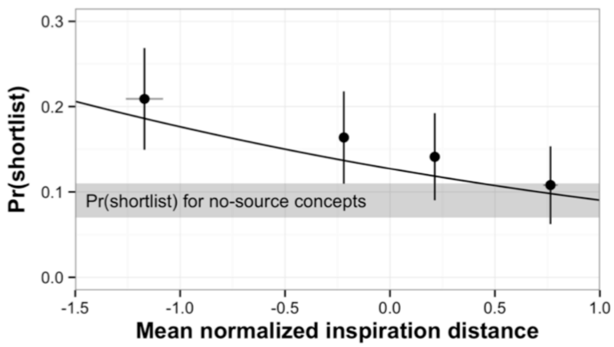

Figure 3: Pr(shortlist) as a function of mean normalized distance. Observed points are mean values from 4 equal *N* bins. Vertical error bars are 95% CI for Pr(shortlist); horizontal error bars are 95% CI for mean distance of the bin. Curve is fitted with mean values for feedback and shortlisted sources, with coefficient of mean distance from the challenge-nesting model. Horizontal gray bar width indicates 95% CI for Pr(shortlist) for concepts with no sources.

and modeling challenge-variation for mean distance gives an essentially zero estimate (u_{1j} = 0.04) and no improvement in model fit from a fixed slope model, $\chi^2(2) = 0.15$, $p = .46$, indicating that the negative effect of inspiration distance was robust across challenges. Since the challenge-nesting model gives the lowest deviance and AIC, we select it as our "best-fitting" model: the 95% CI for the effect of mean distance for this model was –.67 to –.08.

Fig. 3 highlights qualitative interpretations of this effect. Since our distance measure is normalized to have a mean of zero, we can interpret the marked decrease in Pr(shortlist) from the 2nd to 3rd bins as the effect of a shift in the *balance* of near vs far sources (i.e., relying on more far vs. near sources). Further, the horizontal gray bar highlights that there is an overall benefit of building on *any* inspirations: concepts with approximately equivalent amounts of feedback (i.e., mean of 8.43), have a predicted Pr(shortlist = .09, 95% CI = [.07 to .11]; using a logistic model, the coefficient for "any citation" (controlling for feedback) is 0.31, 95% CI = [0.01 to 0.62]). However, the convergence of the fitted and observed lines towards the gray bar as mean distance increases suggests that the benefits of building on sources mainly accrue when building mostly on near inspirations.

Discussion

Summary and Interpretation of Findings

To summarize, we found that — contrary to prior theoretical predictions — relying more on far sources was associated with *worse* creative outcomes, measured by Pr(shortlist), controlling for important control variables, such as the amount of feedback received, and the quality of concepts being built upon. Qualitatively, relying *mostly* on far sources (indicated by a very high mean distance) appears to almost negate the benefits of building on inspirations.

Importantly, this effect was robust across challenges, addressing the concerns raised about potential problem variation. It is also noteworthy that modeling author-nesting did not yield information gain that justified the extra parameters (as suggested by the higher AIC relative to the no-nesting model), suggesting that knowing *how* an idea was developed (e.g., amount of feedback, nature of sources) could yield at least as much insight (if not more, as in our context) into the likely creative outcome of the idea than knowing other characteristics about its author (e.g., intelligence).

Caveats

Some caveats should be discussed before addressing the implications of this study. First, the statistical patterns observed here are *conditional*: i.e., *given that* a concept has cited inspirations as sources, mean distance of those inspirations has a negative relationship with Pr(shortlist). Our data are silent on the effects of mean distance for concepts that did not cite sources. However, these concepts were overall of lower quality; thus, it is unlikely that the negative effect of mean distance can be attributed to attrition (e.g., beneficial far inspirations not being observed). Nevertheless, we should be cautious about making inferences about the impact of sources of inspiration that remain unconscious (since sources in this data are explicitly cited and therefore consciously built upon).

Second, some may be concerned that we have not measured novelty here. Conceivably, the benefits of distance may only be best observed for the novelty of ideas, and not necessarily quality, consistent with some recent work (Franke et al., 2013). However, novelty *per se* is not creativity; thus, we contend that to fully understand the effects of distance on *creativity*, we must consider its impacts on both novelty and quality *together* (as our shortlist measure does).

Implications

These caveats notwithstanding, our results provide strong opposition to the theory that creative ideas are most likely to come from far sources. In light of this opposition (which helps strengthen existing opposing findings), we suggest that the theoretical emphasis on associating creative leaps with far sources may be misguided and require revision.

We should be clear that our findings do not imply that *no* creative ideas come from far sources. Creative ideas can come from both near and far sources; indeed, as our data suggest, some highly creative ideas can come from relying almost not at all on far sources. However, our data do suggest that overreliance on far sources may have a negative impact on creative production (perhaps due to cognitive costs, as mentioned in our introduction).

From a broader perspective, the suggestions of beneficial effects of staying relatively close to the problem domain point to the value of iterative, deep search, a mechanism for creative breakthroughs that may be often overlooked but potentially at least as important as singular creative leaps (Chan & Schunn, in press; Dow et al., 2009; Rietzschel et al., 2007; Sawyer, 2009). Overall, cognitive theories of creativity may benefit from recognizing that there may be multiple parallel paths to creative breakthroughs.

Acknowledgments

This research was supported in part by a Mellon Fellowship to the first author. We are grateful to OpenIDEO for giving us access to this data.

References

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, *43*(5), 997-1013.

Blei, D. M., Ng, A. Y., Jordan, M. I., & Lafferty, J. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 993-1022.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365-376.

Chan, J., Fu, K., Schunn, C. D., Cagan, J., Wood, K. L., & 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*, 081004.

Chan, J., & Schunn, C. (in press). The Impact of analogies on creative concept generation: Lessons from an in vivo study in engineering design. *Cognitive Science.*

Chiu, I., & Shu, H. (2012). Investigating effects of oppositely related semantic stimuli on design concept creativity. *Journal of Engineering Design*, *23*(4), 271-296.

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.

Dow, S. P., Heddleston, K., & Klemmer, S. R. (2009). The efficacy of prototyping under time constraints. In *Proceedings of the 7th ACM conference on creativity and cognition.*

Eckert, C., & Stacey, M. (1998). Fortune favours only the prepared mind: Why sources of inspiration are essential for continuing creativity. *Creativity and Innovation Management*, *7*(1), 1-12.

Enkel, E., & Gassmann, O. (2010). Creative imitation: Exploring the case of cross-industry innovation. *R & D Management*, *40*(3), 256-270.

Franke, N., Poetz, M. K., & Schreier, M. (2013). Integrating problem solvers from analogous markets in new product ideation. *Management Science* (ahead-of-print).

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.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45-56.

Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2013). Inspiration peak: Exploring the semantic distance between design problem and textual inspirational stimuli. *International Journal of Design Creativity and Innovation*, (ahead-of-print), 1-18.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological review*, *114*(2), 211.

Hargadon, A., & Sutton, R. I. (1997). Technology brokering and innovation in a product development firm. *Administrative Science Quarterly*, *42*(4), 716.

Hender, J. M., Dean, D. L., Rodgers, T. L., & Jay, F. F. (2002). An examination of the impact of stimuli type and GSS structure on creativity: Brainstorming versus non-brainstorming techniques in a GSS environment. *Journal of Management Information Systems*, *18*(4), 59-85.

Holyoak, K. J., & Thagard, P. (1996). *Mental leaps: Analogy in creative thought.* Cambridge, MA.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, *12*(1), 3-11.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, *25*(2-3), 259-284.

Malaga, R. A. (2000). The effect of stimulus modes and associative distance in individual creativity support systems. *Decision Support Systems*, *29*(2), 125-141.

Perkins, D. N. (1983). Novel remote analogies seldom contribute to discovery. *The Journal of Creative Behavior*, *17*(4), 223-239.

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.

Sawyer, R. K. (2012). *Explaining creativity: The science of human innovation* (2nd ed.). New York: Oxford University Press.

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.

Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, *27*(1), 1-40.

Ward, T. B. (1998). Analogical distance and purpose in creative thought: Mental leaps versus mental hops. In K. J. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 221–230). Sofia, Bulgaria.

Weisberg, R. W. (2009). On "out-of-the-box" thinking in creativity. In A. B. Markman & K. L. Wood (Eds.), *Tools for innovation* (pp. 23-47). New York, NY.

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