The Importance of Iteration in Creative Conceptual Combination

Joel Chan[a12]

Christian D. Schunn[b]

[a]LRDC Rm 823

University of Pittsburgh

3939 O'Hara St

Pittsburgh, PA 15260, USA

412-648-1109

joc59@pitt.edu

[b]LRDC Rm 821

University of Pittsburgh

3939 O'Hara St

Pittsburgh, PA 15260, USA

412-624-8807

schunn@pitt.edu

[1] **Corresponding author**
[2] **Present address:** 2504B Newell-Simon Hall, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, 412-268-1615, joelchuc@cs.cmu.edu

Abstract

Theories of creative conceptual combination hypothesize that, to generate highly creative concepts, one should attempt to combine source concepts that are very different from each other. While lab studies show a robust link between far combinations and increased novelty of concepts, empirical evidence that far combinations lead to more creative concepts (i.e., both more novel and of higher quality) is mixed. Drawing on models of the creative process, we frame conceptual combination as a divergent process, and hypothesize that iteration is necessary to convert far combinations into creative concepts. We trace conceptual genealogies of many hundreds of concepts proposed for a dozen different problems on a large-scale Web-based innovation platform, and model the effects of combination distance on creative outcomes of concepts. The results are consistent with our predictions: 1) direct effects of far combinations have a mean zero effect, and 2) indirect effects of far combinations (i.e., building on concepts that themselves build on far combinations) have more consistently positive effects. This pattern of effects is robust across problems on the platform. These findings lend clarity to theories of creative conceptual combination, and highlight the importance of iteration for generating creative concepts.

Keywords: Creativity, problem solving, conceptual combination

1. Introduction

How are creative outcomes produced? Conceptual combination is one strategy that has been examined in some depth. It is deceptively simple and process-free in definition: it involves two or more concepts combined into a new concept. Real-world examples of the products of conceptual combination abound, from "mash-ups" and hip-hop sampling in music, to "fusion" cooking, to compound engineered products (like the Apple iPhone, and component/module reuse in engineering). Lab studies have identified a number of different cognitive processes for combining concepts, including property transfer (transferring properties from "helper" concepts to a head concept, e.g., "pet-bird" = "bird you keep in the house and feed when hungry"), hybridization (interpreting a new concept as a "cross" or "blend" between the constituent concepts, e.g., "saw-scissors" = "dual purpose tool that both cuts and saws"), and relational linking (constituent concepts play distinct roles in a thematic relation, e.g., pet-bird = "bird for grooming pets").

Here, we are particularly interested in how conceptual *combination distance* — the degree of semantic distance between the component concepts — influences the creativity of the produced concepts. Specifically, many theorists and eminent creators (Blasko & Mokwa, 1986; Koestler, 1964; Mednick, 1962; Rothenberg, 1979) contend that far combinations are more likely to lead to creative outcomes than near combinations, and numerous anecdotes of eminent creative accomplishments are consistent with this claim (Johansson, 2006; Rothenberg, 1995; Ward, 2001). Is this hypothesis supported by empirical evidence?

Lab studies have consistently shown that *far* combinations — where constituent concepts are semantically distant from each other (e.g., "kitchen utensil" and "bird" vs. "kitchen utensil" and "plate") — lead to more novel combinations (Doboli, Umbarkar, Subramanian, & Doboli, 2014; Gielnik, Frese, Graf, & Kampschulte, 2011; Mobley, Doares, & Mumford, 1992; Nagai, Taura, & Mukai, 2009; Merryl J. Wilkenfeld & Ward, 2001; M. J. Wilkenfeld, 1995; Wisniewski, 1997). A major factor in why this effect occurs is that people generate attributes of the product concept that are emergent, i.e., not characteristic of its constituent concepts. For example, one might say that a “kitchen-utensil bird” is a bird that has a strong jaw for hammering (where neither property is likely to be listed as characteristic of either kitchen utensils or birds when considered separately). Emergent attributes can be generated through first identifying alignable conflicts through analogical mapping (Hampton, 1997) and performing causal reasoning to generate attributes to reconcile those conflicts (Kunda, Miller, & Claire, 1990). Another reason novel concepts are more likely to emerge from combining dissimilar concepts is that people are more likely to think of abstract relations and attributes of constituent concepts (e.g., using metaphor) when those concepts are distantly related (Mumford, Baughman, Maher, Costanza, & Supinski, 1997).

In contrast to the link between combination distance and novelty that has been well established in the lab, the impact of combination distance on idea *creativity* is less clear. Most major models of creativity agree that products are creative if they are both novel and good (of high quality, useful; Boden, 2004; Finke, Ward, & Smith, 1996; Hennessey & Amabile, 2010; Runco, 2004; Sawyer, 2012; Shah, Vargas-Hernandez, & Smith, 2003). However, relatively few studies of conceptual combination and creativity have actually measured quality or creativity. Two lab studies have shown that more distant combinations lead to lower quality ideas (Baughman & Mumford, 1995; Mobley et al., 1992), while one lab study has shown that it has no significant effect, but trending towards higher quality (Doboli et al., 2014). Thus, the connection to quality is unclear. Four lab studies have examined effects on creativity (i.e., the joining of novelty and quality): two found positive effects (Howard-Jones, Blakemore, Samuel, Summers, & Claxton, 2005; Zeng, Proctor, & Salvendy, 2011), while the other two found no effect (Jang, 2014; Siangliulue, Arnold, Gajos, & Dow, 2015), with Siangliulue et al (2015) showing a trend in favor of lower diversity leading to higher creativity.

The relatively small number of studies with mixed results leaves us with uncertainty about the relationship between concept similarity in conceptual combination and creativity. One interpretation of these mixed findings is that far combinations lead only to increased novelty *per se*, not necessarily increased creativity. A related controversy exists in the literature on analogical distance, where studies are divided on whether the most creative analogically inspired ideas come from analogies outside of the problem domain (in other words, from far analogies). Some researchers argue that the best interpretation of the data is that there is no clear/general advantage of far analogies for creative ideation (e.g., Chan, Dow, & Schunn, 2015; Dunbar, 1997; Perkins, 1983; Weisberg, 2009, 2011). Is a similar conclusion (combination distance does not influence creativity) warranted based on the extant empirical data on combination distance? We believe it is plausible, but argue that alternative theoretical interpretations should first be ruled out before accepting it. In this paper, we develop and test one theoretically motivated alternative explanation for the conflicting findings: the benefits of combination distance depend on how much convergence has happened from the point of combination. We argue that, to detect the benefits of combination distance, we need to observe and evaluate the resulting solution path further down its path of development (vs. early on in its development).

To develop our alternative explanation, we draw on a generally shared process model of creativity as involving first, divergent (generating new ideas), then convergent (selecting and building on the best ideas) processes (Amabile, 1983; Finke et al., 1996; Sawyer, 2012;

Simonton, 2011; Wallas, 1926; Warr & O'Neill, 2005). For example, Amabile's (1983) prominent process model prominently includes a movement from divergent processes (response generation) to convergent processes (response validation). Similarly, the Geneplore model (Finke et al., 1996) specifies a Generate phase (initial generation of candidate ideas) followed by an Explore phase (extensive exploration of those ideas). Simplistically, one can view the creative process as linearly progressing from a divergent to a convergent phase. Realistically, creators often go through many divergent-convergent cycles when developing creative products (Herring, Jones, & Bailey, 2009; Jin & Chusilp, 2006). They also sometimes interleave divergent and convergent processes throughout, but transition from earlier periods with more divergence to later periods with less divergence (Atman et al., 2007; Ball, Evans, Dennis, & Ormerod, 1997; Goel & Pirolli, 1992; Shih, Venolia, & Olson, 2011), where convergence on a few promising prototypes becomes necessary to move forward. Overall, there is theoretical consensus that divergent and convergent processes are distinct and jointly necessary for successful creative production, and the creative process moves from an emphasis on divergent processes early on to convergent processes later on.

This theoretical framework provides a principled justification for the hypothesis that far combinations should lead to more creative ideas. If creativity is the production of artifacts that are both new and valuable, then at least some novelty is necessary to create new value. It follows, then, that a creative process that lacked divergence entirely (e.g., only selected from existing ideas) would be highly unlikely to produce a creative idea. Relatedly, models of firm innovation often focus on the tradeoff between exploring uncertain new opportunities and exploiting existing/old certainties (March, 1991). In such models, an exclusive focus on exploitation might be beneficial in the short run, but usually leads to an eventual loss of competitive advantage in dynamically competitive environments. We claim that *far* conceptual combinations in particular — given the usual nature of their conceptual products — are a primarily divergent process for generating new ideas. Therefore, incorporating them into the creative process should eventually increase the likelihood of a highly creative idea, even if they only raise the novelty of ideas considered (but hold quality constant). By contrast, near conceptual combinations could serve both divergent and convergent thinking purposes.

Importantly, understanding far conceptual combination as primarily a divergent process can help explain the conflicting findings on far combinations and creative outcomes. Within this framing, we can draw on the literature on divergent/convergent creative processes to suggest multiple reasons why combination distance might not have an immediate benefit for creativity. First, some researchers argue that a good divergent process increases quality variance in order to make it more likely that the best ideas will be generated (Girotra, Terwiesch, & Ulrich, 2010; Terwiesch & Ulrich, 2009). Therefore, far combination will likely produce both good and bad ideas. Some form of selection process should then be necessary to separate the good ideas from the bad ideas. Secondly, if we conceive of a solution space for creative problems as possessing no more than a few "peaks" (i.e., really good ideas), then statistically there should be many more mediocre or bad ideas than good ideas. It follows from this sparse quality peaks perspective that initial forays into very new regions of the space, if they are "blind" (Simonton, 2011, 2012), will more likely land on mediocre or bad ideas than good ones on the first try. Thus, some time must be allowed to pass in order for some convergent process to select and refine the "good novel" ideas (i.e., to move from the low quality initial landing spot in a novel conceptual region to the nearby high quality variants in that conceptual region). Finally, models and studies of idea generation consistently find that better ideas overall (i.e., combinations of both novelty and quality) tend to be generated later down a solution path (Basadur & Thompson, 1986; Benedek & Neubauer, 2013; Kohn, Paulus, & Choi, 2011; Krynicki, 2014; Nijstad, De Dreu, Rietzschel, & Baas, 2010; Nijstad & Stroebe, 2006; Parnes, 1961; Parnes & Meadow, 1959; Paulus, Kohn, Arditti, & Korde, 2013; Rhodes & Turvey, 2007; Rietzschel, Nijstad, & Stroebe, 2007).

These theoretical insights suggest a potential resolution to the mixed findings regarding combination distance and idea creativity: to observe the benefits of combination distance, one needs to examine its effects well into the convergent phase of the creative process. Given the high-quality-variance nature of far conceptual combination as a creative strategy, a longer convergent phase (i.e., with iteration) may be necessary to convert initially highly novel and highly variable quality ideas to creative solutions (high on both novelty and quality). Therefore, we predict that, if we separately observe the creativity of *direct* and *indirect* conceptual descendants of far combinations, we will see a positive effect for indirect, but not direct, descendants.

We should be clear that we are not making the trivial claim that raw ideas must first be elaborated and iterated on to produce a creative final product (i.e., a main effect of time on idea creativity). What we claim is that initial raw ideas from far combinations in particular are likely to have high novelty, but uncertain utility, and that iterations on these raw ideas (not prototype testing) is necessary to get to creative raw ideas (i.e., concepts that are both highly novel and have high potential utility). In other words, we are making an interaction prediction: the benefits of far over near combinations on idea creativity will only emerge at later time points. Anecdotal accounts of creative discovery by conceptual combination do not clearly specify whether dissimilar conceptual combinations lead to immediately creative raw ideas: some speak of it purely in terms of generating "fresh" (i.e., novel) ideas, being agnostic about the expected potential utility of ideas. Others simply claim that dissimilar conceptual combinations directly yield more creative raw ideas with above average probability. For example, Ward's (2001) analysis of the origins of science fiction author Stephen Donaldson's award-winning *The Chronicles of Thomas Covenant the Unbeliever* fantasy series involves no specification of iteration between the initial leprosy-unbelief conceptual combination and the final idea that formed the overarching theme of the series.

In this paper, we empirically test our theoretically-driven prediction that observations of the benefits of conceptual combination distance vary with the genealogical lag between source and target ideas. Taking a genealogical approach, we trace lagged effects of conceptual combination distance on the creativity of direct and indirect conceptual descendants on a real-world innovation platform. We also address some key methodological issues in prior studies (to increase our confidence in our theoretically motivated hypothesis testing). First, all prior studies examined relatively few creativity problems, and were conducted only in the lab (and therefore under somewhat artificial conditions, and often with toy problems). Some key studies (Doboli et al., 2014; Jang, 2014; Zeng et al., 2011) also had relatively low Ns, making null effects ambiguous and raising potential concerns about effect sizes being significantly overestimated (Button et al., 2013), or even incorrectly estimated as positive/negative in sign (Gelman & Weakliem, 2009). Studies are needed that examine 1) a range of problems with 2) large Ns, 3) under more realistic conditions, and 4) use creativity rather than just novelty or quality in isolation to estimate the general effect of combination distance on creativity. Therefore, the research reported in this paper is conducted on a diverse range of problems, with large numbers of participants working on real world challenges, testing hypotheses with respect to a creative outcome measure that combines both novelty and quality.

2. METHODS

2.1. Dataset

We examine the relationship between combination distance and creative outcomes in the context of OpenIDEO (www.openideo.com), a Web-based innovation platform that addresses a range of social and environmental problems (e.g., managing e-waste, increasing accessibility in elections; see Appendix A for more details on the diverse set of problems sampled for our study). Expert designers from IDEO — a design consulting firm renowned for its creativity — guide platform contributors through a structured design process to produce solutions for these problems that are ultimately implemented for real-world impact ("Impact Stories," n.d.). We focus our analysis in this study on processes and outcomes in three crucial early phases in the process.

- First, in the *inspiration* phase (~1.5 to 4 weeks), contributors help to define the problem space and identify promising solution paths by posting *inspirations*: descriptions of solutions to analogous problems and case studies of stakeholders.
- In the *concepting* phase that follows (for the next 2 to 6 weeks), contributors post

Inspiration: Battery collection. Learn from collection initiatives of old batteries. Consumer can bring old batteries to supermarkets. There collecting containers are installed. Batteries can be recycled and do not pollute the domestic waste. Since people go to the supermarket anyways the hurdle to overcome is quite low.

The organization GRS was founded in 1998 and is installed in Germany. During 2011 44,7 % of the sold batteries where collected after the usage for recycling via the system. You can find more information at http://www.grs-batterien.de/ and in English: http://www.grs-batterien.de/home.html. GRS is the largest take-back system in Europe.

Concept: Rewire.com. Everyone's eager to get the next gadget and gizmo these days, but it's just more money every year to keep up with the latest and greatest stuff. All the while, more than 100 million smartphones and mobile devices lie idle.

Why not take those old smartphones people are willing to donate or throw away and "rewire" them for new uses? An old iPhone is now a media hub for your living room or car. Your old Android HTC is now the remote control for your house lights, security system, and thermostat. Your old tablet is now a digital picture frame on the wall, shuffling through family photos. But smartphones and mobile devices are just part of it - there are so many other electronics to reuse. Rewire's website will offer itself as a collaborative platform where innovative ideas for reusing electronics can get the traction they need to get to the marketplace.

Figure 1. Example inspiration (left) and concept (right) for an OpenIDEO problem about managing electronic waste.

concepts: specific proposed solutions to the stated problem. Crucially, contributors cite concepts or inspirations that serve as sources of inspiration for their idea: this provides our process data for conceptual combination. Figure 1 shows an example inspiration and concept for an OpenIDEO problem about managing electronic waste. They are representative of the typical length and level of specification of inspirations and concepts on the platform.

- In the *shortlist* phase, a subset of concepts for each problem is shortlisted by an expert panel (composed of the OpenIDEO designers and a set of domain experts/stakeholders) for further refinement, based on their creative potential.
- In later stages, these concepts are refined and evaluated in more detail, and then a subset of them is selected for implementation.

We focus on the first three stages given our focus on the use of conceptual combination for generating creative *raw* ideas (the later stages involve many other design processes, such as prototyping). For more details on the dataset, see Chan (2014).

2.2. Sample

The full dataset for this study consists of 2,341 concepts and 4,557 inspirations posted for 12 distinct problems by 1,190 unique contributors. All inspirations and concepts were downloaded (with administrator permission), and a HTML parser was used to extract the following metadata:

1) Concept/inspiration author (who posted the concept/inspiration)
2) Number of comments
3) Shortlist status (yes/no)
4) List of cited sources of inspiration
5) Full-text of concept/inspiration

The current study was conducted with two subsamples of this larger dataset. Specifically, our analysis focused on concepts that (for subsample 1) *directly* cited at least 2 inspirations or (for subsample 2) *indirectly* cited at least 2 inspirations. We define in the next section how we operationalize indirect citations. Our sampling criteria reflect our focus on measuring the effects of conceptual combination distance, which is not measurable with fewer than two sources. The first subsample includes 456 concepts posted by 239 contributors, collectively citing 2,167 unique inspirations. The second subsample includes 522 concepts posted by 281 authors, collectively citing 2,556 unique inspirations.

We were able to obtain professional expertise information (e.g., personal websites, online portfolios, profile pages on company names) posted in the public OpenIDEO profiles of 90 contributors (approximately 1/3 of the authors in the sub-samples). In this sub-sample of the contributors, at least 1/3 are professionals in design-related disciplines (e.g., user experience/interaction design, communication design, architecture, product/industrial design, entrepreneurs and social innovators, etc.) and/or domain experts or stakeholders (e.g., urban development researcher contributing to the vibrant-cities challenge, education policy researcher contributing to the youth-employment challenge, medical professional contributing to the bone-marrow challenge). Thus, from a contributor perspective, our sample includes a range of creative/design expertise, from novice to expert.

2.3. Measures

2.3.1. Conceptual Genealogies

To examine the effects of indirect conceptual descendants, we constructed conceptual genealogies for all concepts in the sample. These genealogies were constructed via breadth-first search through the citation graph gathered in initial data collection: this search first returned all sources that a concept built upon, and then returned all sources that each of these sources built upon (whether they were concepts or inspirations), traversing the conceptual tree to its endpoint. If duplicate entries were encountered, that source was credited at its first appearance in the graph: for instance, if an inspiration I was a direct source for a concept C (at level 1), and also for another concept/inspiration at level 2, it would only be counted once as a level 1 source for C.

In this study, we defined indirect conceptual descendants as inspirations from levels 2 to 4 of each concept's genealogy (see Figure 2): this range choice reflects our goal of examining the effects of sources that are "just recent enough" to have discernible effects (we may not be able to Notice from Figure 2 that indirect sources would also include inspirations cited by cited concepts (i.e., the sources of concepts that acted as immediate sources for the root concept). One way to think about this relationship of the root concept with these indirect sources of other concepts is that (at least part of) the insights/information/ideas contained in those inspirations are "passed on" to the root concept through their incorporation into the concepts immediately cited by the root concept.

2.3.2. Creativity of Concepts

Concept creativity is operationalized as whether a concept was shortlisted. In OpenIDEO, concepts are selected for the shortlist by a panel of expert judges, including the original stakeholders who posted the problem and a team of OpenIDEO designers. Both groups of judges have significant expertise that qualifies them to judge the concepts' creativity: the stakeholders have spent significant time searching for and learning about existing approaches, and the OpenIDEO designers, in addition to their expertise in the general domain of creative design, have spent considerable time upfront with the stakeholders, learning about and defining the problem space.

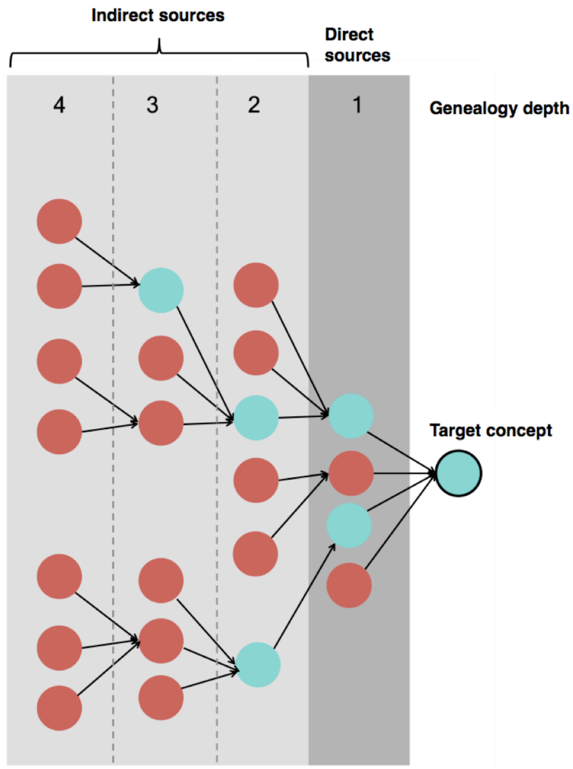

Figure 2. Illustrated example of “indirect” sources as sources in levels 2 to 4 of a specific target concept’s genealogy. Teal circles denote concepts; maroon circles denote inspirations. Note that some indirect sources in this example serve as direct sources for the earlier concepts in this genealogy.

An expert panel is widely considered a “gold standard” for measuring the creativity of ideas (Amabile, 1982; Baer & McKool, 2009; R. T. Brown, 1989; Sawyer, 2012). Further, addressing our need for a creativity measure that jointly considers novelty and quality, we learned from conversations with the OpenIDEO team that the panel’s judgments combine consideration of both novelty and usefulness/appropriateness (here operationalized as potential for impact; A. Jablow, personal communication, May 1, 2014). Additionally, since problems posted on OpenIDEO are unsolved, successful concepts must be different from (and, perhaps more importantly, significantly better than) existing unsatisfactory solutions.

To validate the reported focus of the IDEO panel, we obtained independent external expert judgments on Likert-like scales on separate dimensions of novelty, impact, and feasibility for a subset of concepts in the OpenIDEO data. Specifically, we collected approximately four expert judgments per concept for five of the challenges (N=318 concepts). Reflecting the complex and multidisciplinary nature of the challenges, the expert's ratings had moderate levels of agreement (ICCs of .46, .57, and .63 for novelty, impact, and feasibility, respectively). All three ratings were positively associated with short-list status (novelty $r_{pb} = 0.09$, $p = .10$; impact $r_{pb} = 0.21$, $p = .00$; novelty $r_{pb} = 0.17$, $p = .00$). Fitting a simple logistic regression of shortlist on the three dimensions shows that impact is a strong predictor (b = .51, p = .02). Feasibility is marginally predictive (b = .33, p = .07), while novelty has a positive but nonsignificant estimate (b = .11, p = .63), allaying potential concerns about panel bias against novel concepts.

2.3.3. Combination Distance

A standard approach to measuring combination distance would be to obtain pairwise conceptual distance judgments between all conceptual descendants. However, the scale of the present study (more than 2,000 inspirations would require more than 2 *million* pairwise comparisons) presents formidable challenges to measuring distance using human judgments, from not only a cost standpoint, but also an effectiveness perspective, since the quality of human judgments can deteriorate severely if the workload is too high.

We therefore took a computational approach to measuring combination distance. Specifically, we employed Latent Dirichlet Allocation (LDA; Blei, Ng, Jordan, & Lafferty, 2003) — an unsupervised machine learning technique for learning topical structures from unstructured texts — to learn the semantic space of ideas posted on OpenIDEO, and used that space to estimate semantic distance between inspirations. LDA treats documents as mixtures of latent "topics" (occurring with different "weights" in the mixture), and uses Bayesian statistical learning algorithms to infer the latent topical structure of the corpus (and the topical mixtures for each document) from the co-occurrence patterns of words across documents. With this inferred topical structure, we can then derive conceptual similarity between any pair of documents by computing the cosine between their topical mixtures (which are represented as vectors of topic-weights). Essentially, documents that share dominant topics in similar relative proportions (e.g., primarily about recycling and electronics) are the most similar. This similarity is measured by computing the cosine between their topic-weight vectors, yielding a similarity score between 0 and 1 (where values closer to 1 indicate greater similarity).

LDA has been successfully used to model semantics in many other contexts, including modeling semantic memory representation phenomena (Griffiths, Steyvers, & Tenenbaum, 2007), supporting knowledge discovery and information retrieval in repositories of scientific papers (Griffiths & Steyvers, 2004), and analyzing topical dynamics in social media use (Schwartz et al., 2013). Our application of LDA in the OpenIDEO corpus was validated by examining correlations with human judgments on two sub-samples in the corpus. We collected Likert-scale pairwise similarity judgments for inspirations from 3 research assistants for the first sub-sample, and 5 from the second. Inter-rater agreement was acceptable, aggregate consistency intraclass correlation coefficient (ICC(2,3)) = .46 for the first sub-sample, and ICC(2,5) = .74 for the second sub-sample. The correlation of the LDA cosines with the mean human-judged pairwise similarities was high in both sub-samples, at $r = .54$ and $r = .51$, respectively. Notably, these agreement levels were better than the highest correlation between individual human raters in both subsamples ($r = .39$ and $r = .48$, respectively), reinforcing the value of automatic coding methods for this difficult task. Further details on our implementation and validation of LDA are available in Chan et al. (2014).

Combination distance (hereafter denoted *COMB-DIST*$_{DIR}$ for direct sources, and *COMB-DIST*$_{IND}$ for indirect sources) was measured for each concept as the mean of the reversed pairwise cosines between inspirations cited by that concept (i.e., subtracting from 0, to derive distance rather than similarity). Figure 3 shows an example near and far *COMB-DIST*$_{DIR}$ from the data.

Note that conceptual combination research has tended to focus on pairs of concepts being combined, but here there were often more than just two concepts being combined, especially in

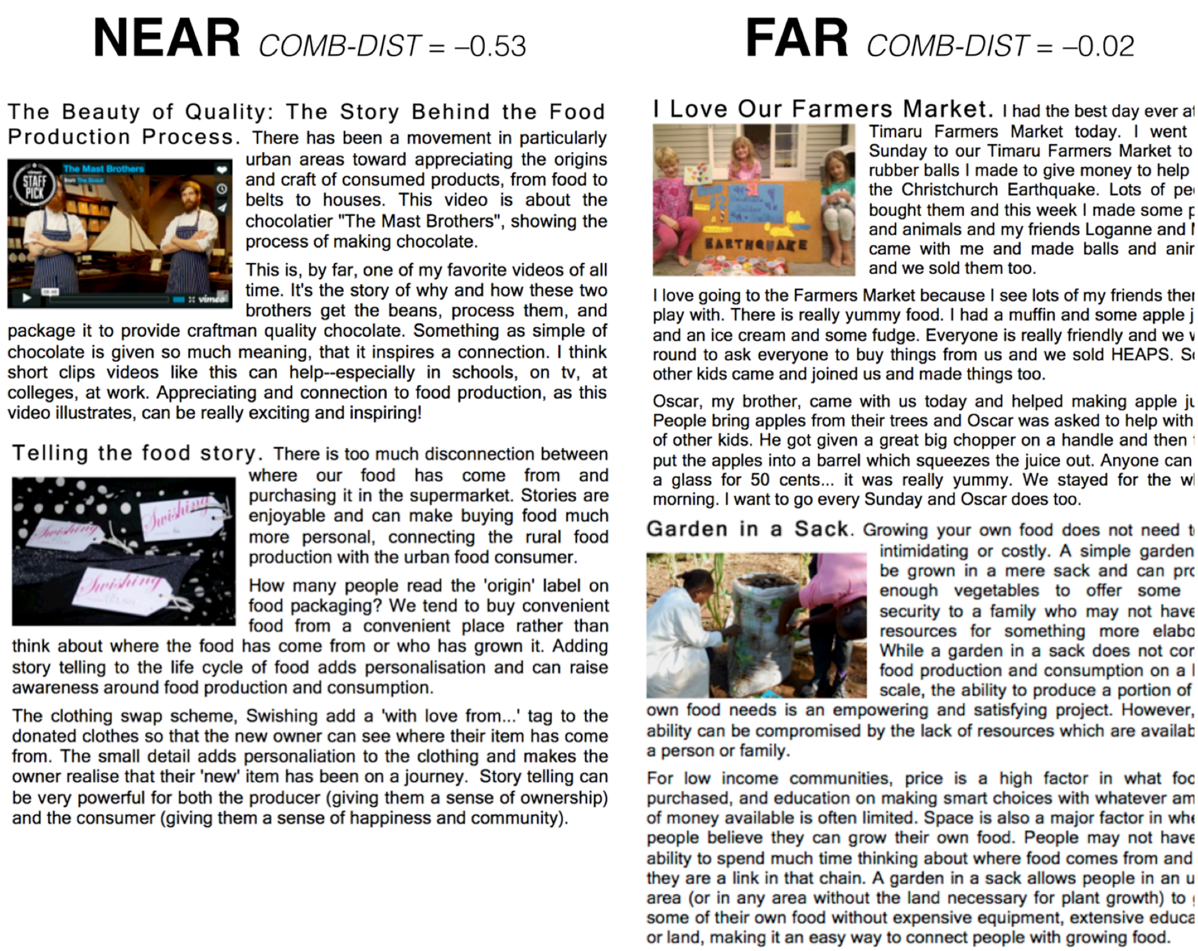

Figure 3. Example near (left) and far (right) combinations according to *COMB-DIST*$_{DIR}$.

the indirect source set. Real world problems are complex, and often involve many subproblems, such that diverse sources must be brought together to solve each of the subproblems. We used mean pairwise distance (the most natural conceptualization of combination distance) to characterize the general diversity among the sources. But it could be that problem solvers brought the most similar pieces together in pairs, or were most influenced by the maximum distance among sources. Therefore we also explored using min and max distance measures rather than mean distance in our analyses. These approaches produced similar results (with slightly more statistical noise), suggesting the patterns we found were not due to idiosyncrasies of how we conceptualized combination distance. However, the added noise when considering min and max distance also suggested our initial intuitions that mean distance is the most appropriate operationalization of combination distance were correct. Therefore, we only report the mean distance results.

2.3.4. Control Measures

To improve our estimates of the effects of combination distance *per se* in this multi-faceted naturalistic dataset, we measured and accounted for other important factors that may influence concept creativity (i.e., we statistically controlled for likely confounds).

Feedback. Feedback can be an important contributor to the quality of a concept. Feedback can provide encouragement, raise issues/questions, or provide specific suggestions for improvement, all potentially significantly enhancing the quality of the concept. Further, feedback may be an alternate pathway to success via combination distance, in that concepts that build on far combinations may attract more attention and therefore higher levels of feedback, which then improve the quality of the concept. On OpenIDEO, concepts receive feedback in the form of comments. We operationalize feedback (labeled here as *FEEDBACK*) as the number of comments received by a given concept.

Quality of cited sources. Concepts that build on existing high-quality concepts (e.g., those that end up being shortlisted) may be more likely to be shortlisted: contributors might incorporate lessons learned from the mistakes and shortcomings, good ideas, and feedback in these high-quality concepts. We operationalize source quality (*SOURCEQUAL*) as the number of shortlisted concepts a given concept builds upon.

Conceptual Source Distance from Problem. Finally, building on a prior study in the OpenIDEO context that showed a positive effect of sources that were conceptually closer to the *problem* domain (Chan et al., 2015), we also control for the distance of sources from the problem domain. Source distance (here labeled *SP-DIST*) is measured for each concept by taking the mean of the reverse cosines between cited inspirations and the problem.

2.4. Analytic Approach

Our goal is to model the creative outcomes of concepts posted by contributors for 12 different problems as a function of combination distance, controlling for other factors. However, contributors are not cleanly nested within problems, nor vice versa; concepts are cross-classified within both authors and challenges (see Figure 4). This cross-classified structure violates assumptions of uniform independence between concepts: concepts posted by the same contributor or within the same problem are likely to be correlated with each other on various dimensions, most importantly overall quality. Failing to account for this non-independence could lead to overestimates of the statistical significance of model estimates (i.e., make unwarranted claims of statistically significant effects). This issue is exacerbated when testing for small effects within large datasets. Additionally, while we are primarily interested in concept-level outcomes, we need to model between-contributor effects to separate out contributor-effects (e.g.,

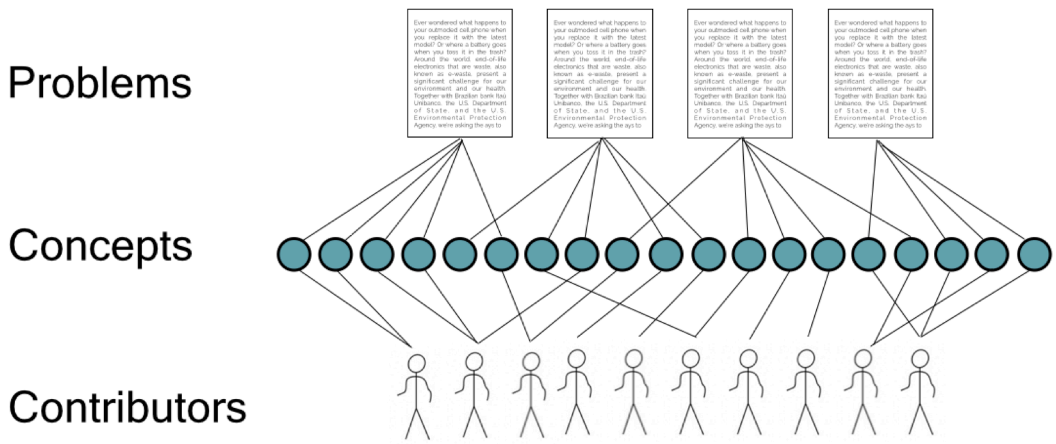

Figure 4. Illustrated cross-classified structure of data

higher/lower creativity, effort) from the impact of sources on individual concepts. Therefore, we employ generalized linear mixed models (GLMMs) to model both fixed effects (of our independent and control variables) and random effects (potential variation of the outcome variable attributable to contributor- or problem-variation and also potential between-problem variation in the effect of combination distance) on shortlist status (a binary variable, which requires logistic, rather than linear, regression).

The following is the general structure of these models (in mixed model notation):

$$\eta_{i(contributorjchallengek)} = \gamma_{00} + \sum_{q} \gamma_{q0} X_{qi} + u_{0authorj} + u_{0challengek}$$

where

- $\eta_{i(contributorjchallengek)}$ is the predicted log odds of being shortlisted for the i^{th} concept posted by the j^{th} contributor in the k^{th} challenge
- γ_{00} is the grand mean log odds for all concepts
- γ_{q0} is a vector of q predictors (q = 0 for our null model)
- $u_{0contributorj}$ and $u_{0challengek}$ model between-contributor and between-challenge variability in mean γ_{00}

We fit our GLMMs using the `glmer` function in the `lme4` package (Bates, Maechler, Bolker, & Walker, 2013) in R (R Core Team, 2013), using full maximum likelihood estimation by the Laplace approximation.

Our general modeling strategy is as follows. First, we fit a reduced model with crossed random effects of challenge and contributor, and fixed effects only of our control measures (i.e., feedback, source quality, and source problem-distance). Because these are theoretically motivated predictors, we leave them in the model regardless of statistical significance. This reduced model serves as a more realistic baseline than the null model; we compare the reduced model to a second *(fixed-slope)* model with the added fixed effect of combination distance. Finally, we fit a third (*random-slope)* model with an added parameter $u_{1challengek}$ to model potential challenge-level random effects on the mean effect of combination distance. To select our final model, we choose the model that meets three criteria: 1) significantly reduces deviance from the null model (low standard for explanatory power), 2) significantly reduces deviance compared to the reduced model from the previous step (higher standard for explanatory power), and 3) has a lower Akaike Information Criterion (AIC) than the previous step to avoid overfitting.

3. Results

3.1. Direct Effects of Combination Distance

We first examine the hypothesis that combining diverse sources leads directly to ideas that are more creative. Recall that this analysis is with the sub-sample of 456 concepts.

3.1.1. Descriptive Statistics

Table 1 summarizes the descriptive statistics and intercorrelations between the variables. There are statistically significant positive correlations between the control variables and Pr(shortlist);

Table 1: Descriptive statistics and intercorrelations between variables for *COMB-DIST*$_{DIR}$

	Descriptives	Correlations			
Variable	M *(SD)*	*FEEDBACK*	*SOURCE SHORT*	*SP-DIST*	*COMB-DIST*$_{DIR}$
SHORTLIST	0.16 *(0.36)*	0.33***	0.11**	–0.10*	–0.01
FEEDBACK	9.14 *(9.92)*		0.12**	0.02	0.05
SOURCESHORT	0.61 *(1.07)*			–0.05	0.10*
SP-DIST	–0.13 *(0.62)*				0.29***
COMB-DIST$_{DIR}$	2.02 *(1.25)*				–

^{m} $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

hence the importance of including them in the models. There are no strong inter-correlations between the predictor variables, alleviating potential concerns about multicollinearity; a variance inflation analysis also shows that having *COMB-DIST*$_{DIR}$ and *SP-DIST* in the same model should not introduce multicollinearity, with variance inflation factors of 1.16 for both variables.

3.1.2. Statistical Models

We fit a series of generalized linear mixed models using full maximum likelihood estimation by the Laplace approximation, with concepts cross-classified within both contributors and problems. We rescale *COMB-DIST*$_{DIR}$ (multiplying it by 10) for easier interpretation (a more meaningful "1-unit" change).

Table 2 presents the model estimates and fit statistics for the GLMMs. The first model is a baseline model fitted with the control variables as predictors. This model yields a large and statistically significant reduction in deviance compared to the null model, χ^2 (2) = 64.70, p = 0.00. Adding a fixed slope for *COMB-DIST*$_{DIR}$ to this model does not provide any meaningful reduction in deviance, with the likelihood ratio being essentially zero, χ^2 (1) = 0.00, p = 0.92, and an increase in the AIC.

Table 2: Model estimates and fit statistics for cross-classified multilevel logistic regressions of Pr(shortlist) on *COMB-DIST$_{DIR}$*, with comparison to baseline model (controls)

	Baseline model (controls)	*COMB-DIST$_{DIR}$*, fixed slope	*COMB-DIST$_{DIR}$* random slope
Fixed effects			
γ_{00}, intercept	–3.08 $^{[-3.37, -2.12]}$	–3.05 $^{[-3.99, -2.12]}$	–3.03 $^{[-4.11, -1.95]}$
γ_{10}, *FEEDBACK*	0.10*** $^{[0.07, 0.12]}$	0.10*** $^{[0.07, 0.13]}$	0.10*** $^{[0.07, 0.13]}$
γ_{20}, *SOURCESHORT*	0.25^{m} $^{[-0.10, 0.35]}$	0.25^{m} $^{[-0.03, 0.52]}$	0.26^{m} $^{[-0.03, 0.54]}$
γ_{30}, *SP-DIST*	–0.49^{m} $^{[-0.71, 0.10]}$	–0.50^{m} $^{[-1.05, 0.04]}$	–0.54* $^{[-1.08, -0.00]}$
γ_{40}, *COMB-DIST$_{DIR}$*		0.01 $^{[-0.27, 0.30]}$	0.03 $^{[-0.28, 0.33]}$
Random effects			
$u_{0authorj}$	0.47	0.47	0.44
$u_{0challengek}$	0.71	0.71	1.63
$u_{1challengek}$			0.05
Model fit statistics[1]			
Deviance	323.57	323.57	321.74
AIC	335.57	337.57	339.74

^{m} $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$; 95% CI (Wald) = $^{[lower, upper]}$

The point estimate for the effect of a change of .10 in *COMB-DIST$_{DIR}$* (remember that *it* is rescaled in this model) is also essentially zero (see Figure 5), albeit with a fairly wide confidence interval. To ensure that this wide confidence interval is not due to one or two outlier problems overwhelming an overall positive or negative trend across the problems, we estimate an additional model with a random slope for *COMB-DIST$_{DIR}$*. Visually inspecting the posterior modes for the slope of *COMB-DIST$_{DIR}$* for each problem (see Figure 5), we see an even scatter about the mean value, with 6 challenges having a positive sign for the coefficient for diversity, and 6 challenges having a negative sign. A binomial sign test with a null hypothesis of equal probability for positive and negative effects of diversity estimates a two-tailed *p*-value of 1.00 of observing either 6 or fewer positive or 6 or more negative signs in 12 "trials" (although this

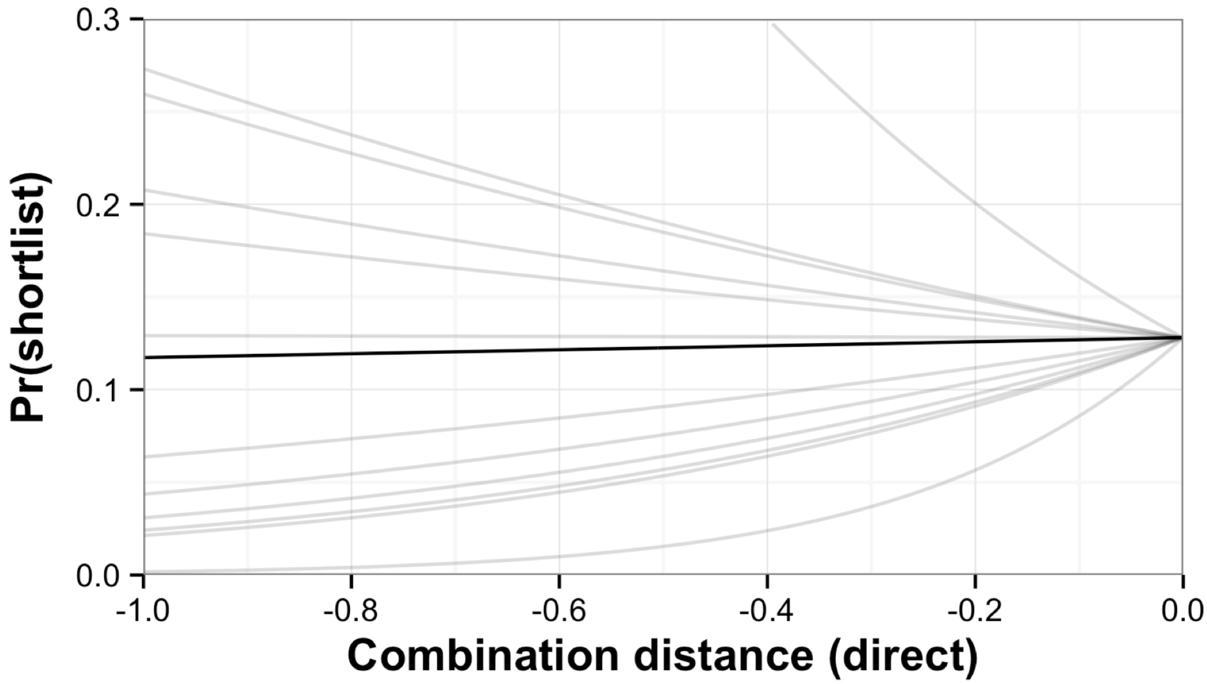

Figure 5. Model-fitted relationship between combination distance and Pr(shortlist). Fitted values evaluated at mean values of *FEEDBACK*, *SOURCEQUAL*, and *SP-DIST*. Greyed lines are fitted from posterior mode estimates of the slope of $COMB\text{-}DIST_{DIR}$ for each problem.

binomial test outcome should be intuitively obvious to the reader, we present it here to parallel the same test conducted on the indirect problem-specific slopes).

The graph in Figure 5 gives the impression that there are clear effects of combination distance, which can be positive or negative depending upon the problem. However, it is important to note that there is not strong evidence in support of significant problem moderation: the model estimates low problem-variance, does not meaningfully decrease variance from the fixed slope model, χ^2 (2) = 1.83, p = .23 (p-value is halved, heeding common warnings that a likelihood ratio test discriminating two models that differ on only one variance component may be overly conservative, e.g., Pinheiro & Bates, 2000), and also further increases AIC, calling into question whether the estimated problem variation is meaningful. In other words, the most parsimonious interpretation given these data is that direct combination distance has no effect (i.e., we select the baseline model as our final model), although it is possible that an even larger dataset would find problem-specific effects. Most importantly, these data argue strongly against a general benefit of combination distance of direct sources on idea creativity.

3.2. Indirect Effects of Conceptual Diversity

We now turn to the analysis of the effects of combination distance of indirect sources. Recall that this analysis is with the sub-sample of 522 concepts.

3.2.1. Descriptive Statistics

Descriptive statistics and bivariate correlations are given in Table 4 and **Error! Reference source not found.**. There are no strong correlations among the predictors, giving little cause for concerns about multicollinearity.

3.2.2. Statistical Models

As before, we estimate a series of generalized linear mixed models to analyze the relationship between *COMB-DIST*$_{IND}$ and Pr(shortlist). We exclude *COMB-DIST*$_{DIR}$ from our models for a number of reasons. First, it does not add predictive value (as we saw in the preceding analysis). Second, including it as predictor would exclude concepts that had indirect

Table 3: Descriptive statistics for indirect combination distance measures

Variable	Valid N	Min	Max	Mean	Median	SD
SHORTLIST	522	0	1	0.15	0	0.36
FEEDBACK	522	0	67	9.01	6	10.02
SOURCESHORT	522	0	11	0.67	0	1.06
SP-DIST	522	-2.93	1.67	-0.11	-0.01	0.73
COMB-DIST$_{IND}$	522	-0.73	-0.02	-0.18	-0.14	0.10

Table 4: Bivariate correlations for indirect combination distance measures

Variable	*FEEDBACK*	*SOURCE SHORT*	*SP-DIST*	$COMB\text{-}DIST_{IND}$
SHORTLIST	0.34***	0.13**	–0.11*	0.04
FEEDBACK		0.11*	–0.01	0.13**
SOURCESHORT			–0.05	0.19***
SP-DIST				–0.02

^{m} $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

inspirations as sources, but had less than two direct inspiration sources, reducing our N to 381 for no predictive gain. Finally, our estimates of the effects of $COMB\text{-}DIST_{IND}$ do not change with $COMB\text{-}DIST_{DIR}$ in the model.

The model estimates are given in Table 5. As before, we begin with a baseline controls model, which gives a large and statistically significant reduction in deviance compared to the null model, χ^2 (3) = 63.70, p = 0.00. In contrast to $COMB\text{-}DIST_{DIR}$, adding a fixed slope for $COMB\text{-}DIST_{IND}$ to the baseline model yields a marginally significant reduction in deviance, χ^2 (1) = 3.26, p = 0.07, and a decrease in AIC, mitigating concerns about overfitting.
The model estimates that a .10 change in $COMB\text{-}DIST_{IND}$ corresponds to an increase of approximately .45 in the log-odds of being shortlisted (see Figure 6). Holding all the other predictors at their mean values, changing from a $COMB\text{-}DIST_{IND}$ of –0.20 (close to the mean value in the sample) to –0.10) increases Pr(shortlist) from 0.13 to 0.19. Again, the CI for the effect is relatively wide. However, in contrast to *COMB-DIST,* the estimated positive effect of $COMB\text{-}DIST_{IND}$ did not appear to vary by problem. An additional model with a random slope for $COMB\text{-}DIST_{IND}$ estimates very low problem-variance, does not meaningfully decrease variance from the fixed slope model, χ^2 (2) = 0.26, p = .44 (p-value is halved), and also further increases AIC. Therefore, we select the fixed slope model as our final model for this analysis.

Table 5: Model estimates and fit statistics for cross-classified multilevel logistic regressions of Pr(shortlist) on $COMB\text{-}DIST_{IND}$, with comparison to baseline model (controls only)

	Baseline model (controls)	With $COMB\text{-}DIST_{IND}$, fixed slope	With $COMB\text{-}DIST_{IND}$ random slope
Fixed effects			
γ_{00}, intercept	$-2.80^{[-3.44, -2.16]}$	$-1.98^{[-3.10, -0.86]}$	$-2.12^{[-3.10, -0.86]}$
γ_{10}, *FEEDBACK*	$0.09^{***\,[0.06, 0.12]}$	$0.09^{***\,[0.07, 0.12]}$	$0.09^{***\,[0.07, 0.12]}$
γ_{20}, *SOURCESHORT*	$0.16^{[-0.08, 0.39]}$	$0.12^{[-0.12, 0.35]}$	$0.12^{[-0.12, 0.35]}$
γ_{30}, *SP-DIST*	$-0.44^{*\,[-0.82, -0.07]}$	$-0.45^{*\,[-0.83, -0.06]}$	$-0.45^{*\,[-0.83, -0.06]}$
γ_{40}, $COMB\text{-}DIST_{IND}$		$0.45^{m\,[-0.04, 0.94]}$	$0.34^{m\,[-0.04, 0.94]}$
Random effects			
$u_{0authorj}$	0.12	0.13	0.12
$u_{0challengek}$	0.60	0.88	1.35
$u_{1challengek}$			0.03
Model fit statistics			
Deviance	372.65	369.39	369.13
AIC	384.65	383.39	387.13

^{m} $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$; 95% CI (Wald) = $^{[lower, upper]}$

Importantly, when estimating the posterior modes for the effect of diversity for each problem, we see that none of the 12 challenges has a negative estimate (see Figure 6). A binomial sign test with a null hypothesis of equal probability for positive and negative effects of diversity estimates a two-tailed p-value of 0.0005 of observing either 0 or fewer positive or 12 or more negative signs in 12 “trials”. In other words, the effect of combination distance of indirect sources is consistent across problems in a way that is very unlikely to have arisen by chance.

3.3. Iteration Chain Depth and Earliness of Inspirations

One plausible alternative explanation for our findings is that concepts that cite indirect sources with high combination distance are better not because of their combination distance of those

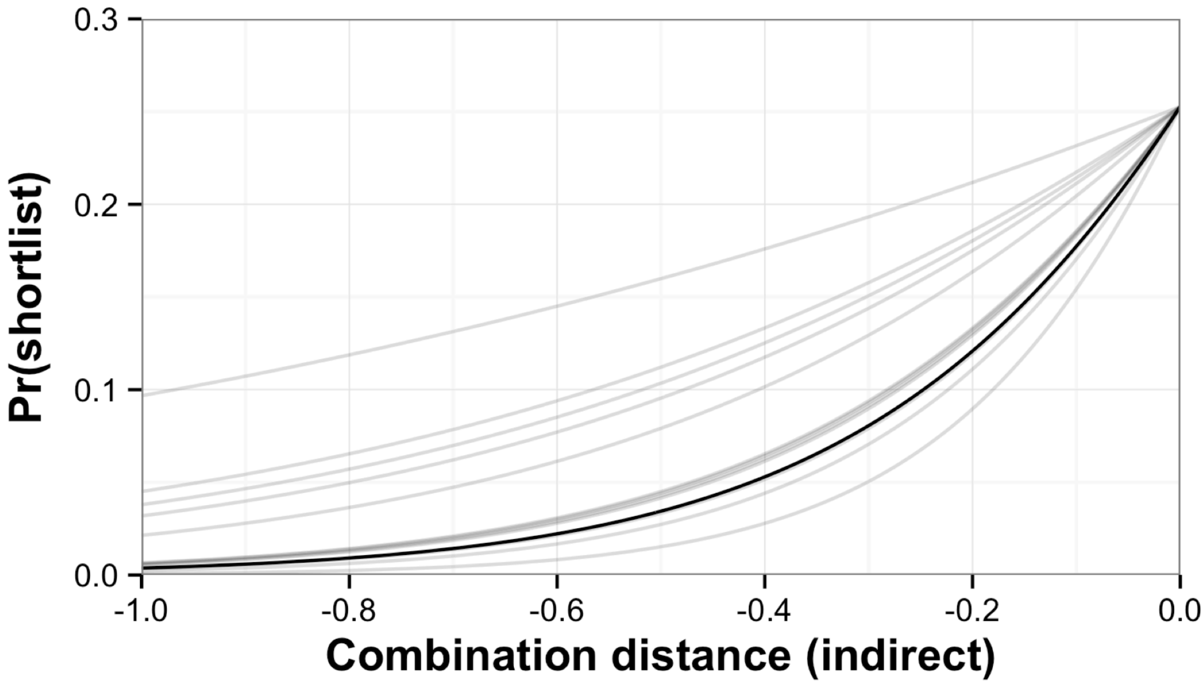

Figure 6. Model-fitted relationship between combination distance of indirect sources and Pr(shortlist). Fitted values evaluated at mean values of *FEEDBACK*, *SOURCEQUAL*, and *SP-DIST*. Greyed lines are fitted from posterior mode estimates of the slope of $COMB\text{-}DIST_{IND}$ for each problem.

sources, but because those sources were inspirations posted earlier in the challenge. These earlier inspirations might lead to better concepts for a variety of reasons unrelated to combination distance, e.g., they might be of higher quality because they are posted by "early adopters" to the challenge who are more motivated, or they might articulate base concepts more clearly (e.g., because they have more time for iteration).

Depth in iteration chain (our main variable of interest) in fact had a small positive association with earliness in time (i.e., when the inspiration was posted). At the inspiration level, inspirations that were posted earlier are slightly more likely to show up deeper in iteration chains, r = 0.13, p < .001. Inspirations that are cited as both immediate and indirect sources (*M* =

9.5 days into the challenge, *SE* = 0.16) tend to be posted earlier than inspirations that are only ever cited as immediate sources (*M* = 12.2 days, *SE* = 0.49).

However, the mean earliness of concepts' cited inspirations is not predictive of their creative success. Estimating a generalized linear mixed model with feedback, source quality, and mean inspiration earliness as fixed effects and challenge and contributor as random effects, we find that the estimated effect of earliness is near-zero, b = 0.03, 95% CI = [–0.01, 0.08], p = 0.17. This model does not significantly improve fit over the reduced control variables model (i.e., with just fixed effects of feedback, source quality, and challenge and contributor random effects), LRT χ^2 (1) = 1.76, p = 0.18 (AIC = 716.86 vs. 716.62). Similarly, mean depth of cited inspirations in a chain is not predictive of concepts' creative success. Adding mean chain depth to the reduced control variables model does not significantly improve fit, LRT χ^2 (1) = 1.46, p = 0.23 (AIC = 717.17 vs. 716.62), and the model estimates no reliable effect of mean chain depth, b = –0.16, 95% CI = [–0.42, 0.09], p = 0.22. Therefore, mere earliness of cited inspirations cannot explain the interaction between iteration depth and the effect of inspiration diversity on creative outcomes.

4. Discussion

4.1. Summary

Our goal in this paper was to examine conceptual combination as a strategy for generating creative ideas. Theories of conceptual combination and creativity suggested the hypothesis that distant conceptual combinations are especially likely to lead to highly creative ideas, but the past empirical support for this hypothesis has been uneven. Drawing on broader theories of the creative process (Amabile, 1983; Finke et al., 1996; Sawyer, 2012; Simonton, 2011; Wallas, 1926; Warr & O'Neill, 2005), we formulated a theoretical framework that situates distant conceptual combination within the creative process as a *divergent* creative strategy. This theoretical framework yielded a novel refinement of the distant combination hypothesis: the benefits of conceptual combination distance are more likely to be seen with a genealogical lag between source and target ideas. In other words, we predicted that it takes time for distant combinations to yield their creative fruits.

The current study's findings provided empirical support for this refined hypothesis. As predicted, analyzing combination distance of indirect sources indeed yielded different results than direct sources. Specifically, we found that the mean effect of direct combination distance, though slightly trending in a positive direction, was essentially zero (with some potential problem variation). In contrast, combination distance of *indirect* sources was a positive predictor of creative outcomes, and this effect was robust across problems. Thus, distant combinations do appear to be especially likely to lead to highly creative outcomes, but only if they are "indirect" (i.e., sources of one's sources). Importantly, we also demonstrate that the contrast between direct and indirect sources is not explained by the mere earliness of the indirect sources.

4.2. Strengths and Limitations

Before we draw out the larger implications of this study, we first note some strengths and limitations of this study. First, we note that we were able to strike a favorable tradeoff between external validity (real designers solving real creative problems) and statistical power, which is rare in creativity research (i.e., typical studies of real designers have smaller Ns than lab studies, not larger Ns, as in the current study). This feature narrows the gap considerably from the current findings to generalizations in real-world creative cognition (Dunbar & Blanchette, 2001; Henrich, Heine, & Norenzayan, 2010). Another strength of our study is our creative outcome measure, which combines both novelty and quality, and follows the gold standard expert panel approach. This allows us to think more holistically about the effects of combination distance on *creativity*, not just novelty and/or quality in isolation (a major gap in prior work).

One limitation of the current work is that our correlational study design does not allow us to make strong causal claims. Relatedly, because our data source was preexisting naturalistic behavioral traces online, we were not able to precisely isolate cognitive processes at a fine-grained level, as one might be able to in the laboratory, or obtain psychological control measures (e.g., participants' familiarity with inspirations). These are legitimate concerns, and to some degree are inherent tradeoffs of an *in vivo* vs. *in vitro* approach (Dunbar & Blanchette, 2001). However, three features of our study mitigate concerns about spurious statistical associations. First, our findings align well with prior theory and laboratory findings on the potential benefits of combination distance. Indeed, the external validity from the *in vivo* approach strengthens our confidence that the laboratory findings generalize to real-world creative cognition. Second, unlike some other observational designs, our study does include a temporal asymmetry between the predictor and outcome variables (we know that sources were built upon *before* shortlisting), which is a notable indicator of causal direction. Finally, our statistical analysis accounts for problem variation, contributor effects, and a variety of other important control variables, mitigating concerns over endogeneity. Nevertheless, future randomized experiments are necessary to fully establish causality.

Additionally, some might be concerned that the unique context of the study — e.g., Web-based context, focus on socially relevant problems — might limit generalizability to other creative problems. This is a legitimate concern, given the recent controversy in the literature over the extent to which creative processes are domain-general or domain-specific (Plucker, Beghetto, Sternberg, Grigorenko, & Singer, 2004; Simonton, 2009). We have reason to believe that our results should generalize well to creativity in design-related domains (e.g., engineering/product/architectural design) given the nature and diversity of problems in our sample. Solutions to these problems drew on a wide range of domain knowledge (e.g., public policy, human-computer-interaction, social and decision sciences, education) and likely involved reasoning over multiple levels of systems (e.g., individual decision-making, communities). Thus, the cognitive processes and knowledge involved in generating concepts in our study are likely to have significant overlap with other design-related domains. Nevertheless, we encourage further studies that explore how these findings might generalize to (or be different in) other forms of creative thought, such as artistic creativity and scientific discovery.

Similarly, because of the public, Web-based context of the study, contributors were probably not posting every idea they had, regardless of quality. This is a different dynamic from typical lab studies (e.g., of brainstorming) where participants are asked to write down all ideas that come to mind. It is likely that in a less filtered context, the moderating effects of iteration might be much more pronounced (but still be fundamentally similar): our findings suggest that 2 to 4 iterations on a concept combined from distant concepts are sufficient to make it creative, but more iterations may be required in more realistic settings where less self-filtering is going on.

4.3. Conclusions

Returning to the overall question of the nature of human creative cognition, our findings align with other studies and arguments that have highlighted the importance of iteration, broadly construed, in the creative process (Chan & Schunn, 2015; Dow, Heddleston, & Klemmer, 2009; Mecca & Mumford, 2013; Nijstad et al., 2010; Rietzschel et al., 2007; Weisberg, 2011; Merryl J. Wilkenfeld & Ward, 2001; Yu & Nickerson, 2011). This emerging body of work suggests that iteration provides pathways to not only higher quality ideas (Dow et al., 2009; Yu & Nickerson,

2011) but also more novel ideas (Nijstad et al., 2010; Rietzschel et al., 2007), and is an advanced application of cognitive strategies like analogy (Chan & Schunn, 2015). Here, we add to this body of work, showing that iteration is a critical partner process in creative conceptual combination. Again, we emphasize that this line of research concerns the importance of iteration for the development of *ideas*, not final products. The emerging picture is that good ideas rarely come in singular creative leaps, fully formed like Athena from the brow of Zeus; instead, they more often come from the sweat of his brow building on the labors of others.

Our findings are consistent with the predominant view of the optimal temporal ordering of divergence and convergence (iteration) in the creativity literature. Most theories of creativity posit a “flare and focus” view of creativity: diverge first, then converge. Similarly, many authors and practitioners warn of the dangers of premature solution selection (e.g., T. Brown & Katz, 2009; Vogel, Cagan, & Boatwright, 2005). Berg (2014) showed in a series of elegant experiments that ideas that build first on very novel ideas that are then “infused” with more well-trodden idea components end up with a more optimal combination of novelty and quality than ideas that begin with well-trodden ideas and then are “infused” with novelty.

While there is consensus in the literature on the optimal order of divergence and convergence, one might wonder about the optimal *balance* between divergence and convergence. In the present work, we observed a monotonically positive relationship between indirect combination distance and creativity. Insofar as far combinations increase divergence (and “variation” in the variation-selection view of things), this relationship makes sense. However, this monotonically positive relationship stands in contrast with some other work in the context of group brainstorming that has found no overall benefit of increased divergence (through nominal groups) for the creativity of the final idea that is selected (e.g., Rietzschel, Nijstad, & Stroebe, 2010). It might be fruitful to examine whether different divergent strategies (e.g., brainstorming, distant conceptual combinations) have different trajectories (e.g., non-monotonic, polynomial) in terms of their effects on the final idea. A genealogical methodological approach (as exemplified in the current study) might be helpful for exploring this new space of questions.

5. Acknowledgements

This research was supported in part by a National Science Foundation #1360013 Doctoral Dissertation Improvement Grant to the first author.

6. References

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

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, *45*(2), 357.

Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of Engineering Education*, *96*, 359–379.

Baer, J., & McKool, S. S. (2009). Assessing creativity using the consensual assessment technique. In C. S. Schreiner (Ed.), *Handbook of Research on Assessment Technologies, Methods, and Applications in Higher Education* (pp. 65–77). Hershey, PA.

Ball, L. J., Evans, J. S. B. ., Dennis, I., & Ormerod, T. C. (1997). Problem-solving Strategies and Expertise in Engineering Design. *Thinking and Reasoning*, *3*(4), 247–270. http://doi.org/10.1080/135467897394284

Basadur, M., & Thompson, R. (1986). Usefulness of the Ideation Principle of Extended Effort in Real World Professional and Managerial Creative Problem Solving*. *The Journal of Creative Behavior*, *20*(1), 23–34. http://doi.org/10.1002/j.2162-6057.1986.tb00414.x

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5.* Retrieved from http://CRAN.R-project.org/package=lme4

Baughman, W. A., & Mumford, M. D. (1995). Process-analytic models of creative capacities: Operations influencing the combination-and-reorganization process. *Creativity Research Journal*, *8*(1), 37–62.

Benedek, M., & Neubauer, A. C. (2013). Revisiting Mednick's Model on Creativity-Related Differences in Associative Hierarchies. Evidence for a Common Path to Uncommon Thought. *The Journal of Creative Behavior*, *47*(4), 273–289. http://doi.org/10.1002/jocb.35

Berg, J. M. (2014). The primal mark: How the beginning shapes the end in the development of creative ideas. *Organizational Behavior and Human Decision Processes*, *125*(1), 1–17. http://doi.org/10.1016/j.obhdp.2014.06.001

Blasko, V. J., & Mokwa, M. P. (1986). Creativity in advertising: A Janusian perspective. *Journal of Advertising*, *15*(4), 43–72.

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

Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. New York.

Brown, R. T. (1989). Creativity: What are we to measure? In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of Creativity* (pp. 3–32). New York NY.

Brown, T., & Katz, B. (2009). *Change by design : how design thinking transforms organizations and inspires innovation*. New York: Harper Business.

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. http://doi.org/10.1038/nrn3475

Chan, J. (2014). *The impact of sources of inspiration on the genesis of creative ideas* (Doctoral Dissertation). University of Pittsburgh. Retrieved from http://d-scholarship.pitt.edu/22743/

Chan, J., Dow, S. P., & Schunn, C. D. (2015). Do The Best Design Ideas (Really) Come From Conceptually Distant Sources Of Inspiration? *Design Studies*, *36*, 31–58.

Chan, J., & Schunn, C. (2015). The impact of analogies on creative concept generation: Lessons from an in vivo study in engineering design. *Cognitive Science*, *39*(1), 126–155.

Doboli, A., Umbarkar, A., Subramanian, V., & Doboli, S. (2014). Two experimental studies on creative concept combinations in modular design of electronic embedded systems. *Design Studies*, *35*(1), 80–109. http://doi.org/10.1016/j.destud.2013.10.002

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*.

Dunbar, K. N. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 461–493). Washington D.C.

Dunbar, K. N., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: the case of analogy. *Trends in Cognitive Sciences*, *5*(8), 334–339.

Finke, R. A., Ward, T. B., & Smith, S. M. (1996). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.

Gelman, A., & Weakliem, D. (2009). Of Beauty, Sex and Power. *American Scientist*, *97*, 310–316.

Gielnik, M. M., Frese, M., Graf, J. M., & Kampschulte, A. (2011). Creativity in the opportunity identification process and the moderating effect of diversity of information. *Journal of Business Venturing*, *27*(5), 559–576. http://doi.org/10.1016/j.jbusvent.2011.10.003

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea generation and the quality of the best idea. *Management Science*, *56*, 591–605.

Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science*, *16*, 395–429.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proc Natl Acad Sci U S A*, *101 Suppl 1*, 5228–35. http://doi.org/10.1073/pnas.0307752101

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychol Rev*, *114*(2), 211–44. http://doi.org/10.1037/0033-295X.114.2.211

Hampton, J. A. (1997). Emergent attributes in combined concepts. In T. B. Ward, S. M. Smith, & J. Viad (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 83–110). Washington, DC: American Psychological Association.

Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology*, *61*, 569–98.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 1–23.

Herring, S. R., Jones, B. R., & Bailey, B. P. (2009). Idea generation techniques among creative professionals. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on* (pp. 1–10). IEEE.

Howard-Jones, P. A., Blakemore, S.-J.-. J., Samuel, E. A., Summers, I. R., & Claxton, G. (2005). Semantic divergence and creative story generation: An fMRI investigation. *Cognitive Brain Research*, *25*(1), 240–250.

Impact Stories. (n.d.). Retrieved from http://www.openideo.com/content/impact-stories

Jang, S. (2014). The Effect of Image Stimulus on Conceptual Combination in the Design Idea Generation Process. *Archives of Design Research*, *112*(4), 59. http://doi.org/10.15187/adr.2014.11.112.4.59

Jin, Y., & Chusilp, P. (2006). Study of mental iteration in different design situations. *Design Studies*, *27*(1), 25–55.

Johansson, F. (2006). *The Medici effect: What elephants & epidemics can teach us about innovation*. Boston, MA: Harvard Business Press.

Koestler, A. (1964). *The act of creation*. Oxford, England: Macmillan.

Kohn, N. W., Paulus, P. B., & Choi, Y. (2011). Building on the ideas of others: An examination of the idea combination process. *Journal of Experimental Social Psychology*, *47*(3), 554–561.

Krynicki, F. (2014). *Methods and models for quantitative analysis of crowd brainstorming*. University of Waterloo.

Kunda, Z., Miller, D. T., & Claire, T. (1990). Combining Social Concepts: The Role of Causal Reasoning. *Cognitive Science*, *14*(4), 551–577. http://doi.org/10.1207/s15516709cog1404_3

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87.

Mecca, J. T., & Mumford, M. D. (2013). Imitation and Creativity: Beneficial Effects of Propulsion Strategies and Specificity. *The Journal of Creative Behavior*. http://doi.org/10.1002/jocb.49

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220–232.

Mobley, M. I., Doares, L. M., & Mumford, M. D. (1992). Process analytic models of creative capacities: Evidence for the combination and reorganization process. *Creativity Research Journal*, *5*(2), 125–155.

Mumford, M. D., Baughman, W. A., Maher, M. A., Costanza, D. P., & Supinski, E. P. (1997). Process-Based Measures of Creative Problem-Solving Skills: IV. Category Combination. *Creativity Research Journal*, 59.

Nagai, Y., Taura, T., & Mukai, F. (2009). Concept blending and dissimilarity: factors for creative concept generation process. *Design Studies*, *30*(6), 648–675.

Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *EUROPEAN REVIEW OF SOCIAL PSYCHOLOGY*, *21*, 34–77. http://doi.org/10.1080/10463281003765323

Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: a cognitive model of idea generation in groups. *Pers Soc Psychol Rev*, *10*(3), 186–213. http://doi.org/10.1207/s15327957pspr1003_1

Parnes, S. J. (1961). Effects of extended effort in creative problem solving. *Journal of Educational Psychology*, *52*(3), 117–122.

Parnes, S. J., & Meadow, A. (1959). Effects of “brainstorming” instructions on creative problem solving by trained and untrained subjects. *Journal of Educational Psychology*, *50*(4), 171–176.

Paulus, P. B., Kohn, N. W., Arditti, L. E., & Korde, R. M. (2013). Understanding the Group Size Effect in Electronic Brainstorming. *Small Group Research*, *44*(3), 332–352. http://doi.org/10.1177/1046496413479674

Perkins, D. N. (1983). Novel Remote Analogies Seldom Contribute to Discovery. *The Journal of Creative Behavior*, *17*(4), 223–239.

Pinheiro, J. C., & Bates, D. M. (2000). *Linear mixed-effects models: basic concepts and examples*. Springer.

Plucker, J. A., Beghetto, R. A., Sternberg, R. J., Grigorenko, E. L., & Singer, J. L. (2004). Why Creativity Is Domain General, Why It Looks Domain Specific, and Why the Distinction Does Not Matter. In *Creativity: From potential to realization* (pp. 153–167). Washington, DC, US: American Psychological Association.

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rhodes, T., & Turvey, M. T. (2007). Human memory retrieval as Lévy foraging. *Physica A: Statistical Mechanics and Its Applications*, *385*(1), 255–260.

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.

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2010). The selection of creative ideas after individual idea generation: Choosing between creativity and impact. *British Journal of Psychology*, *101*(1), 47–68. http://doi.org/10.1348/000712609X414204

Rothenberg, A. (1979). *The emerging goddess: The creative process in art, science, and other fields*. University of Chicago Press Chicago.

Rothenberg, A. (1995). Creative cognitive processes in Kekule's discovery of the structure of the benzene molecule. *The American Journal of Psychology*, 419–438.

Runco, M. A. (2004). Creativity. *Annual Review of Psychology*, *55*, 657–687.

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

Schwartz, A. H., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., … Ungar, L. H. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLoS ONE*, *8*(9), e73791. http://doi.org/10.1371/journal.pone.0073791

Shah, J. J., Vargas-Hernandez, N., & Smith, S. M. (2003). Metrics for measuring ideation effectiveness. *Design Studies*, *24*(2), 111–134.

Shih, P. C., Venolia, G., & Olson, G. M. (2011). Brainstorming Under Constraints: Why Software Developers Brainstorm in Groups. In *Proceedings of the 25th BCS Conference on Human-Computer Interaction* (pp. 74–83). Swinton, UK, UK: British Computer Society. Retrieved from http://dl.acm.org/citation.cfm?id=2305316.2305331

Siangliulue, K., Arnold, K. C., Gajos, K. Z., & Dow, S. P. (2015). Toward Collaborative Ideation at Scale: Leveraging Ideas from Others to Generate More Creative and Diverse Ideas. In *Proceedings of CSCW'15*.

Simonton, D. K. (2009). Varieties of (Scientific) Creativity: A Hierarchical Model of Domain-Specific Disposition, Development, and Achievement. *Perspectives on Psychological Science*, *4*(5), 441–452.

Simonton, D. K. (2011). Creativity and discovery as blind variation: Campbell's (1960) BVSR model after the half-century mark. *Review of General Psychology*, *15*(2).

Simonton, D. K. (2012). Creativity, Problem Solving, and Solution Set Sightedness: Radically Reformulating BVSR. *The Journal of Creative Behavior*, *46*(1), 48–65.

Terwiesch, C., & Ulrich, K. T. (2009). *Innovation tournaments: Creating and selecting exceptional opportunities*. Boston, MA: Harvard Business Press.

Vogel, C. M., Cagan, J., & Boatwright, P. B. H. (2005). *The Design of Things to Come: How Ordinary People Create Extraordinary Products*. Upper Saddle River, NJ.

Wallas, G. (1926). *The art of thought*. New York, NY.

Ward, T. B. (2001). Creative cognition, conceptual combination, and the creative writing of Stephen R. Donaldson. *The American Psychologist*, *56*(4), 350–354.

Warr, A., & O'Neill, E. (2005). Understanding Design As a Social Creative Process. In *Proceedings of the 5th Conference on Creativity & Cognition* (pp. 118–127). New York, NY, USA: ACM. http://doi.org/10.1145/1056224.1056242

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.

Weisberg, R. W. (2011). Frank Lloyd Wright's Fallingwater: A Case Study in Inside-the-Box Creativity. *Creativity Research Journal*, *23*(4), 296–312. http://doi.org/10.1080/10400419.2011.621814

Wilkenfeld, M. J. (1995). *Conceptual combinations: Does similarity predict emergence?*

Wilkenfeld, M. J., & Ward, T. B. (2001). Similarity and Emergence in Conceptual Combination. *Journal of Memory and Language*, *45*(1), 21–38.

Wisniewski, E. J. (1997). Conceptual combination: Possibilities and esthetics. In T. B. Ward, S. M. Smith, & J. Viad (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 51–81). Washington, DC: American Psychological Association.

Yu, L., & Nickerson, J. V. (2011). Cooks or cobblers? Crowd creativity through combination. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1393–1402). ACM.

Zeng, L., Proctor, R. W., & Salvendy, G. (2011). Fostering creativity in product and service development: validation in the domain of information technology. *Hum Factors*, *53*(3), 245–70.

APPENDIX A: OpenIDEO Data Additional Details

To give a better sense of the range of problems addressed on the platform, the following are the problem titles (as seen by participants) for the 12 problems in the study sample:

1. How might we increase the number of registered bone marrow donors to help save more lives?
2. How might we inspire and enable communities to take more initiative in making their local environments better?
3. How can we manage e-waste & discarded electronics to safeguard human health & protect our environment?
4. How might we better connect food production and consumption?
5. How can technology help people working to uphold human rights in the face of unlawful detention?
6. How might we identify and celebrate businesses that innovate for world benefit and inspire other companies to do the same?
7. How might we use social business to improve health in low-income communities?
8. How might we increase social impact with OpenIDEO over the next year?
9. How might we restore vibrancy in cities and regions facing economic decline?
10. How might we design an accessible election experience for everyone?
11. How might we support web entrepreneurs in launching and growing sustainable global businesses?
12. How can we equip young people with the skills, information and opportunities to succeed in the world of work?