Distributed Analogical Idea Generation: Inventing with Crowds

Lixiu Yu, Aniket Kittur, Robert E. Kraut
Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
{lixiuyu, nkittur, robert.kraut}@cs.cmu.edu

ABSTRACT
Harnessing crowds can be a powerful mechanism for increasing innovation. However, current approaches to crowd innovation rely on large numbers of contributors generating ideas independently in an unstructured way. We introduce a new approach called *distributed analogical idea generation*, which aims to make idea generation more effective and less reliant on chance. Drawing from the literature in cognitive science on analogy and schema induction, our approach decomposes the creative process in a structured way amenable to using crowds. In three experiments we show that distributed analogical idea generation leads to better ideas than example-based approaches, and investigate the conditions under which crowds generate good schemas and ideas. Our results have implications for improving creativity and building systems for distributed crowd innovation.

Author Keywords
Crowdsourcing; schema; innovation; creativity; analogy

ACM Classification Keywords
H.5.3. Group and Organization Interfaces

INTRODUCTION
Innovation is becoming increasingly driven by crowds. Many companies are turning to crowds to provide solutions to a given problem: Innocentive (innocentive.com) posts research challenges for its crowd of scientists and inventors to solve, 99Designs (99designs.com) uses the crowd to generate a variety of graphic designs, and Threadless (threadless.com) manufactures t-shirts from crowd-submitted designs. Other firms, like Quirky (quirky.com), ask the crowd to identify interesting problems as well as solutions to them.

However, crowd innovation can also fail or produce poor quality ideas. For example, Quirky has developed and manufactured 340 distinct products from ideas generated using a crowdsourcing paradigm, but these products have been culled from hundreds of thousands of contributed ideas. Such numbers are indicative of current crowd innovation approaches, which rely on large numbers of individuals or small teams working independently, under the assumption that each has a small chance of contributing a valuable idea or solution. In such an environment enough people may produce a few good ideas, but the process is highly dependent on chance, and results in many ideas that provide little value to the company and many members of the crowd who do not gain any monetary benefits. Improving the effectiveness of crowd idea generation could thus have significant benefits to both companies and crowds.

Researchers have investigated ways to boost the quality of the idea generation process. Research in multiple domains shows that creativity is often built on existing work [e.g., 14], and an established technique is to use multiple previous examples to inspire new ideas [3]. However, while prior examples may help an inventor generate new ideas, they may also stifle creativity [27].

One promising approach to these challenges may be found in studies of analogy in discovery and innovation. Analogy is a powerful and prevalent mechanism for leveraging prior knowledge into new insights [11]. Scientific and technological breakthroughs rarely occur in a vacuum, and are often the result of taking a solution from a source domain and transferring it to a target domain [10, 15]. Studies of creative problem solving show that recognizing similarity in structure between problems and then applying the common structure is an effective way of solving problems [12].

However, previous work in analogy has focused on well-defined problems and well-specified target and source domains [10, 11], and even in these constrained domains problem solvers have tremendous difficulty spontaneously identifying an idea's structure and transferring it to another domain [12]. Furthermore, such approaches still rely on a single inventor or small team to engage in the entire creative process from identifying analogies to generating solutions for the problems.

In this paper we present a novel approach to generating innovative ideas through directed analogical transfer that results in more creative ideas than the common approach of using multiple examples. Furthermore, our approach provides a structured way to decompose the creative idea gen-

CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada
Copyright 2014 ACM 978-1-4503-2473-1/14/04…$15.00.
http://dx.doi.org/10.1145/2556288.2557371

eration process into multiple steps so that it is amenable to using multiple groups of people. By breaking the analogical transfer process up into discrete stages we can recruit different individuals in different phases as well as select and prune ideas. Through a series of studies we show that distributed analogical idea generation leads to better ideas than example-based approaches, and investigate the conditions under which crowds generate good schemas and ideas.

RELATED WORK

Crowd Innovation

Many attempts to crowdsource creative work and idea generation have used online contests, in which an open call is made and a small number of submissions are chosen as winners [24]. Contests generally aim to collect contributions from many contributors, relying on large numbers of submissions to surface at least one quality product. A number of approaches have been proposed to improve the efficiency of the contest approach. For example, some contests encourage submitters to provide feedback to each other [31] or to collaborate in small teams [16]. Our work differs from these because it does not require direct interaction between inventors, which can introduce challenges in coordination, incentives and scaling [21].

Examples and Creativity

One established technique in design practice is leveraging examples of prior work [3]. At IDEO, arguably the most influential design firm in the world, product designers store samples of past products and reuse them when there is valuable connection with current problems [14]. Studies on crowd creativity have found that providing and combining multiple examples can increase the quality of new ideas [33], and combining novel features of examples is especially useful [34]. Presenting examples early and multiple times can be especially effective improved creative work [23].

However, studies have also found that examples can lead to unintentional conformity effects, constraining the generation of creative ideas [27]. For example, a recent study on online design contests found that showing designers high quality design examples caused them to produce designs that were more homogeneous and less distinctive [22]. Perhaps because of the tension between the use of examples to inspire creativity and its tendency to constrain it, there are neither consistent findings about the examples' effects on the quality of new ideas nor a consensus about how to use examples to improve creativity. Whether presenting examples leads to a positive effect is contingent on many factors, such as the number, variety, features, and timing of presented examples. Furthermore, selecting the right examples is a challenge: with so many examples available, inventors could be easily overwhelmed.

Design Patterns

A concept that is related to our current work is that of *design patterns:* standard solutions to common problems that can be used in many different situations [2]. Design patterns have been used in domains including architecture, software engineering, and website design [30]. However, a challenge with design patterns is that it is often difficult to induce them: in architecture and software design, experts develop the design patterns. More scalable computational methods to induce patterns (e.g., [29]) have been largely limited to highly structured domains such as structured websites. Our work aims to support the induction and application of patterns by the crowd, increasing the scalability and generality of the idea generation process.

DISTRIBUTED ANALOGICAL IDEA GENERATION

Many innovative solutions and discoveries occur when people take an idea from one domain and apply it to another domain, a process known as *analogical transfer*. A canonical example used in numerous studies of analogical transfer is the problem of a doctor who wants to destroy a tumor with X-rays but not harm the surrounding tissue [12]. A solution to the problem is to use a divide and converge strategy: divide the x-ray into multiple small beams converging on the tumor. People are more likely to solve this problem if they first see how an analogous problem was solved (e.g., a general who conquers a castle by dividing his army into small groups and having them attack from different directions). By generating a high-level structure or *schema* of the problem and its solution, people can then use that schema to solve other problems that are structurally similar to the original problem, even if they are in different domains and have very different surface features [10, 17, 18, 20].

While a potentially powerful mechanism for innovation, researchers have found analogical transfer difficult to robustly induce. There are a number of essential steps to successful analogical transfer, any of which can cause it to fail. One core challenge is getting people to generate appropriate schemas. A schema refers to a structured mental representation consisting of entities and the relations between them[1]. To induce a schema people must generate a representation of a problem abstracted from the surface features of the source example (i.e., inducing *divide and converge* from the general/castle example) [12]. Even if an appropriate schema has been generated, another challenge has been getting people to notice that the schema is relevant and could be applied to the target domain (i.e., recognizing that *divide and converge* could solve the X-ray problem) [12].

As a result of these challenges, research on analogical transfer in cognitive science and design has focused on solving well-defined problems and using analogies from well-specified target and source domains given to participants [12, 17, 6, 25]. Unfortunately, these constraints have important limitations when applied to creative idea generation and invention, where problems are often not well-defined, target domains are not given, and where the goal may be to find a new problem, a new solution or a new domain – or potentially all three. For example, take the case where an

[1] For example, a schema for a love triangle involves three entities each of whom loves and is loved by different entities.

inventor wants to develop a new idea for a product. Unlike the analogy example described above, the inventor does not have a defined target problem or even a defined target domain; indeed, the goal is to create something novel that has not been previously imagined.

How can we develop a robust, repeatable, distributed process that can promote analogical transfer for creative idea generation? A key development is that inventors now have access to repositories of thousands of proposed ideas. Many of these ideas may share the same schema – for example, Table 1 shows three ideas from Quirky.com that share the same high-level stability schema: stabilizing something by attaching it to a more stable object. The presence of diverse, real-world embodiments of schemas in the form of proposed ideas enables us to introduce a more general framing of the analogical transfer process that supports crowd invention:

Step 1: Choose two products as examples. In this case the inventor chooses Props and Tether, shown in Table 1.

Step 2: Induce a common analogical schema from these examples. Here the inventor may induce a *stability* schema: stabilizing a mobile object by connecting it to a more stationary object. This schema provides a representation of both ideas at a more abstract level that is not specific to the source domain or surface features of the examples.

Step 3: Identify other domains where the schema could be applicable. The inventor may think of domains where things are unstable, such as bumpy bus rides, airplane turbulence, earthquakes or people with Parkinson's disease.

Step 4: Applying the schema to the target domain to generate a new product. For example, applying the *stability* schema to the bus ride domain might result in the Eclipse idea shown in Table 1, which stabilizes a person's head by connecting it via a strap to their seat.

We call this process *distributed analogical idea generation* (illustrated in Figure 1). This process has a number of attributes that makes it especially appropriate for crowdsourcing creative tasks, in which the four steps just described can be distributed among different groups. In contrast to most design contests or online brainstorming, it exploits the wisdom of the crowds for multiple subtasks, including selecting relevant examples, identifying common schemas and generating domains, problems, and solutions. This decomposition helps crowd members know what to do by providing a structured process. It supports contributions in small units of effort by decomposing the creative process into smaller steps. Because it does not require direct interaction among crowd members, it minimizes coordination costs and scales well. And because the crowd can generate multiple possibilities at each step, the process may be directed by selectively retaining intermediate products with the most potential.

In the following three experiments we empirically investigate the value of distributed analogical idea generation and the conditions under which it is most effective. As there are

The stability schema: Stabilize a mobile object by using an attachment device to connect it to a more stable object.	
	Props: This headphone keeper is designed to keep your earbuds tethered to you, making them easily accessible when you want to return to your musical oasis. Worn around the neck, Props extends like arms to your earbuds, keeping them in place and in the vicinity of your ears.
	Tether: Washing your wine glasses can be risky business, especially when you are using your dishwasher. Save your stemware with Tether, a flexible plastic rod that stabilizes your stemware as it goes through the cycle.
	Eclipse: Eclipse is a plush, padded eye mask designed to keep you comfortable while you travel. The mask comes with a webbing strap which can be wrapped around a headrest, and velcros to the back of the eye mask to keep your head from bobbing while sleeping.

Table 1. Examples and the shared schema.

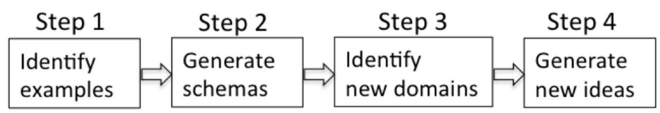

Figure 1: The distributed analogical idea generation.

a number of challenges at each step that could derail the process, we order the presentation of experiments working backwards from the desired final state. Thus in Experiment 1 we investigate whether given an analogical schema participants generate better ideas than if not given such a schema (i.e., Steps 3 & 4). We describe this experiment first as it tests our core hypothesis. In Experiment 2 we then examine the conditions under which participants are able to generate good schemas (Steps 1 & 2). Finally in Experiment 3 we demonstrate how the full process can be accomplished in a distributed manner by using different participants for different steps.

EXPERIMENTS

Experiment 1: Generating Ideas from Analogical Schemas

In this experiment we aim to test whether the crowd generates better ideas from analogical schemas than from examples. We operationalize a specific type of schema in the context of this paper as an idea schema, consisting of a *problem* an idea aims to solve (e.g., preventing something fragile from breaking) and a *mechanism* for doing so (e.g., using a flexible cord to secure it to a stable object). Importantly, a schema abstracts away surface details to focus on the underlying structure [10, 12].

One important factor governing the effectiveness of a schema is likely to be how concrete versus abstract is its representation. Abstract schemas may help people avoid fixating on specific features, promoting far transfer to diverse domains and solutions. However, abstract schemas may be difficult for people to use if they are at too high a level to usefully constrain the possible domains and solutions for an idea. Although there have been conflicting views regarding the role of examples and schemas in analogical transfer, abstract schemas alone seem to be difficult to communicate if they are not grounded in specific examples [26]. However, these examples may lead to greater fixation on details that may be detrimental [27]. For example, problem solvers may become fixated on details like the Velcro or plastic rods used to attach one object to another, constraining their creativity.

To examine the effects of abstract versus concrete representation of schemas, we created two ways of presenting schemas: as abstract schemas alone (*Use schemas*) and as abstract schemas illustrated with concrete examples (*Use schemas and see examples*). We compared these to two control conditions in which participants brainstorm product ideas after merely seeing examples (*See examples*) or being encouraged to use these examples as the basis of their product ideas (*Use examples*).

Participants

In this and all experiments, participants were recruited by posting tasks to Amazon Mechanical Turk [19]. Overall 209 workers participated in Experiment 1. Forty-seven percent of participants were women, and 90% were native English speakers. Their average age was 31 and ranged from 19 to 69.

The stability schema: Stabilize a mobile object by using an attachment device to connect it to a more stable object.	
	Eclipse: Eclipse is a plush, padded eye mask designed to keep you comfortable while you travel. The mask comes with a webbing strap which can be wrapped around a headrest, and velcros to the back of the eye mask to keep your head from bobbing while sleeping.
The separation schema: Detach one object from another by using comb-like mechanisms.	
	Broom Groomer: it is a step-on dustpan with added functionality that makes sweeping easier. Rubber "teeth" on the back of the dustpan let you quickly and easily comb out dust bunnies from your broom's bristles, while a smooth rubber "lip" on the front keeps the dustpan flush with the floor so nothing slips through the cracks.
The space schema: Create space and organize things by raising the work surface and creating storage.	
	Spacebar: The Space Bar is a desk accessory that minimizes clutter while providing additional USB ports for your computer. After a long day of work, simply slide your keyboard into the designated space below the shelf and store your office miscellany-keys, digital camera, etc.-up top.

Table 2. Experiment 1: Examples and schemas.

Design and Procedure

After they accepted the task, participants were randomly assigned to one of the four conditions and asked to generate a product idea. In all conditions, participants were told that their ideas would be evaluated in terms of *practicality* (the product can be designed and produced easily), *usefulness* (the product solves an important problem) and *novelty* (the product hadn't been thought of by others). The conditions were:

1. See examples. Participants brainstormed a new product idea after seeing the three examples shown in Table 2. They were not explicitly required to use any of the examples as the inspiration for their product idea.

2. Use examples. Participants were first asked to read the three examples in Table 2 as well as summaries describing their novel properties. We created these summaries to control for processing and engagement differences involved in the *Use schema and see examples* condition (the third condition). Because participants in the latter condition were asked to read both the examples and the schemas, they may have been more likely to engage in deeper cognitive processing than the naturalistic *See examples* condition. To avoid this engagement difference, which might affect the final results, we asked people to read both the examples and their summaries. Specifically they were told,

"*Product ideas have different features, which are special properties of the products to attract consumers. The above product ideas have their own features that make them different from their similar products. For example,*

Eclipse: A mask that not only keeps the light out, but stabilizes your head by using a strap so it doesn't bob around during travel.

Broom groomer: It makes sweeping easier just by adding rubber "teeth"."

Space bar: It minimizes clutter while providing additional USB ports for your computer.

Participants were asked to generate one new product idea based on the given examples.

3. Use schemas and see examples. Participants were asked to read the three examples shown in Table 2 along with experimenter-generated schemas for each (referred to as 'design rules' in the instructions). These schemas are also shown in Table 2 above the description of each example. Participants were then asked to generate a new product idea by using one of the schemas.

They were told, "*Now please generate a new product idea by using one of the design rules.*"

4. Use schemas. Participants only saw the three schemas shown in Table 2 without seeing the examples. They were asked to generate a new product idea by using one of the schemas.

Rating idea quality

As rating the creativity of ideas is by nature a subjective judgment, we base our outcome measures on prior research that argues that creative ideas need to be both novel and practical [7]. We considered a product novel if it was not obvious and was different from existing products on the market[2]. We operationalized practicality in terms of *usefulness* to the user of the product and the *practicality* of manufacturing it today. For example, a teleportation device might be both novel and useful, but is infeasible to manufacture.

Two judges blind to the experimental condition rated each product idea on three 7-point Likert scales of novelty, usefulness, and practicality. One judge was the first author, and the second judge was an oDesk worker with skills in product development and marketing. After several rounds of training and discussion, the judges achieved high inter-rater reliabilities of .90, .86 and .88 for practicality, usefulness, and novelty respectively, calculated as the Intraclass Correlation Coefficient (ICC) [8]. The final quality of the ideas was computed as the geometric mean of the three dimensions to weight against ideas that were low on any given dimension [9].

Below we present an example of a highly-rated idea generated from the *separation* schema (to detach one object from another by using comb-like mechanisms). This idea was rated high in novelty, usefulness, and practicality.

The problem when kneading dough is that the dough gets stuck in between your fingers. This makes it difficult to clean your hands, and you also lose a lot of the dough in the process. The solution is a device that has a space for each of your fingers. It will come in three sizes so that you can pick the size that best suits the size of your fingers. You run your fingers through the device and the extra dough is removed. The surface of the finger-comb does not allow the dough to stick to it, and thus is easily removed and cleaned. The extra dough can be added to the kneaded dough to [eliminate] waste and speed up clean-up time.

Analysis and Results

The means and standard errors of the ratings for the ideas are shown in Figure 2. We were interested in the quality of ideas produced by different idea generation processes. Among the four conditions, two conditions were analogy-based (*Use schema and see examples* and *Use schema*), and two conditions were example-based (*See examples* and *Use examples*). To determine whether analogy-based idea generation led to better ideas than example-based idea generation we ran a regression analysis with the dependent variable the quality of the ideas, with a planned contrast between the analogy- and example-based conditions. The results showed that analogy-based idea generation resulted in significantly better ideas than the example-based conditions, $b = .32$, $t(208) = 3.34$, $p < .001$; $d = .47$.

[2] Raters judged novelty based on their experience rather than a market survey or patent because we wanted to assess an idea's relative novelty compared to products that most people would know about, rather than using the patent office's standard that the idea was unique in historical time [28, p.169].

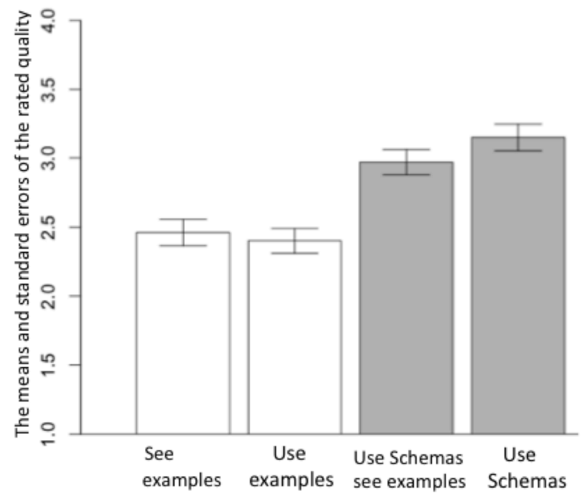

Figure 2. Experiment 1: the quality of the ideas

We further examined the question of whether analogy-based idea generation would be helped or hindered by showing concrete examples along with the analogical schema. However, we found no significant difference between the *Use schemas* and *Use schemas and see examples* conditions, $b=.12$, $t(113)=.94$, $p=.35$. Thus the benefits of using analogical schemas appear to be robust to whether they are shown grounded in concrete examples.

Although prior research has suggested that combining schema and examples improves analogical problem solving [28], our task was different from typical analogical transfer in that we asked people to generate completely new ideas by finding both problems and solutions. In this context it is possible that adding examples to schemas might produce two conflicting effects: fixating people on surface features, while helping them better understand the schemas. These two forces might counteract each other's influence on quality. Further research is needed to understand the conditions under which these results may hold.

Experiment 2: Generating Schemas from Examples

After demonstrating that schemas can improve idea generation, we turn to the question of how to help people generate good analogical schemas. Even though previous research shows that good schemas are critical to analogical transfer, getting people to induce them has been remarkably challenging [13]. Researchers have asked participants to summarize analogous problems, provide a verbal statement of the underlying principle or generate a diagrammatic representation, all with limited success [12, 13]. The most effective method for schema induction appears to involve having participants explicitly map multiple problems sharing a schema onto each other [12]. When people find correspondences between these instances, they are more likely to create good schemas that focus on deep relations rather than surface features. Based on these findings, we predicted that

	Pivot Power: a creative outlet with adjustable outlets. It has a flexible form that bends into circular, semi-circular, and zig-zag shapes to fit around furniture and in tight spaces. Every plug fits into every outlet.

Table 3. The contrasting example used in Experiment 2.

having people induce schemas from multiple examples would help them produce useful schemas.

Additionally, prior research has found that adding contrasting examples can enhance the quality of generated schemas [13]. Showing a contrasting example that does not share an underlying schema could help highlight similarities between the ideas that do share the schema and therefore result in better schema induction.

Participants

Overall 145 workers participated in Experiment 2: 53% percent were women, and 91% were native English speakers. Their average age was 33 and ranged from 19 to 66.

Design and Procedure

All participants in this experiment were exposed to a training session, similar across experimental conditions, where they were asked to generate a schema for the *Broom Groomer* product idea from Table 2. We then showed them a correct schema generated by the experimenters for the example (the *separation* schema in Table 2). Participants were informed that a good schema should meet the following criteria:

"*It specifies a purpose (e.g., detaches things); it specifies a mechanism (e.g., uses comb-like features); it shouldn't be too vague (e.g. "make things easy" is too vague); it shouldn't have too many details (e.g., the details of the given idea, such as the color or the dustpan, should not be included).*"

After training, participants were shown one or more examples from a fixed set of ideas (depending on condition) and asked to generate a single schema that applied to one or more examples. Specifically, the instructions stated:

"*A schema refers to the underlying structure of an idea. It is a description or template for how to solve a problem that can be used in many different situations. Can you find a schema that applies to one or more of the above product ideas?*"

Participants were randomly assigned to one of four conditions. In the *one example, two examples,* and *three examples* conditions, participants respectively saw a random set of one, two or three of the examples from Table 1, all of which share the *stability schema* (though the schema was not provided to them). Additionally, a *contrasting example* condition was used in which participants saw two random examples sharing the stability schema and one product idea ("Pivot Power") that did not share the *stability* schema or any other schema from the examples in Table 3. Thus the *contrasting example* condition had the same total number of examples as the *three examples* condition, and the same number of analogous examples as the *two examples* condition.

Rating schema quality

In order to rate participants' schemas, we developed a rubric for what makes a high quality schema based on previous literature on schema induction and transfer [12] and adapting it to the domain of creative idea generation and product development. We operationalize a good schema as capturing the essential problem and mechanism from examples without including specific example details (e.g., "Detach one object from another by using comb-like mechanisms"). Schemas could be scored poorly for multiple reasons: they could be too general to describe the essence of the idea because they do not describe the problem or mechanism (e.g., "Make sweeping easier" does not include a mechanism); or they could be too specific because they include features from specific examples (e.g., "Add a rubber comb to a dustpan"). Two raters blind to the condition rated participants' schemas using this rubric on a 7-point Likert scale. Schemas from three participants were removed because they ignored or misunderstood the instructions. After several rounds of training and discussion, the judges achieved high inter-rater reliability of .93. The reported results are based upon the averaged scores of the two raters.

Analysis and Results

Table 4 shows the means and standard deviations of all conditions. We first tested the effects of the number of examples on the quality of the generated schemas. We ran a regression analysis with the quality of the schemas as the dependent variable. We performed two contrast coding: a linear contrast testing whether schemas improved with an increasing number of examples, and a curvilinear contrast testing for diminishing returns of the number of examples. The results revealed that the linear relationship between *One example*, *Two examples* and *Three examples* was significant, $b = .72$, $t(113) = 3.69$, $p < .001$ and the curvilinear relationship was not, $b = .12$, $t(113) = .95$, $p = .35$. These results suggest that adding analogous examples facilitates schema generation, and that adding additional analogous examples beyond the first continue to improve the resulting schemas.

We next examined the effects of adding a contrasting example on the quality of the generated schemas. Two conditions had been included to test this: *Three examples* and *Contrasting examples*. The *Three examples* condition had a third similar example and *Contrasting examples* had a third contrasting example. A regression analysis comparing *Three examples*, *Two examples* against *Contrasting examples* conditions showed that the difference between *Three examples*

Condition	Mean	SD	Freq.
1. One example	2.09	1.82	45
2. Two examples	3.16	1.68	31
3. Three examples	3.53	1.78	38
4. Contrasting examples	2.68	1.68	28
Total	2.86	1.74	142

Table 4. Experiment 2: Effects of example number and type on schema quality.

and *Contrasting examples* was significant, $b = .85$, $t(65) = 1.94$, $p < .05$, and the difference between *Two examples* and *Contrasting examples* was not significant, $b=.48$, $t(58) = 1.06$, $p = .29$. These results suggest that adding a contrasting example is not as useful as adding another analogous example. One explanation for this may be that while Gick and Paterson [13] found benefits for adding contrasting examples, their contrasting examples had dimensions in common and could be aligned with the analogous examples. Our contrasting examples on the other hand were very different from the analogous examples, and thus may not have been useful for drawing participants' attention to schema-relevant aspects of the examples.

Experiment 3: Distributed Analogical Idea Generation

In Experiments 1 and 2 we established the building blocks of the distributed analogical idea generation process. We have shown that participants were able to generate better ideas using experimenter-generated schemas than using examples (Experiment 1), and also identified conditions under which participants can generate good schemas themselves (Experiment 2). We now close the loop by examining whether the schemas generated by participants in Experiment 2 also lead to better ideas than using examples. In doing so we aim to answer two questions: first, whether the quality of the schema generated in Experiment 2 affects the quality of the resulting ideas; and second, whether the process can be run end-to-end in a distributed way; that is, using different individuals at different stages.

There are a number of potential advantages to having a distributed idea generation process. Breaking the task into stages enables multiple people to engage in the process iteratively, potentially benefiting from the diversity of their ideas and backgrounds. It reduces the participation costs of the task, opening it up to those who may be less committed or have less expertise. It also enables quality control to take place between stages. For example, once people generate schemas, those schemas can be evaluated and only the good ones used in the idea generation step. In theory, having multiple stages could also allow for flexible control and direction of the ideation process, as the examples, schemas, domains, and ideas could be selectively pruned at each stage to focus on areas of interest to the process creator.

In Experiment 3 we selected some of the good and poor schemas produced in Experiment 2 as input and asked new participants to generate ideas based on them. We compared the quality of the ideas to those produced after seeing good or poor schemas to those produced after seeing examples (using the procedures from the *See examples* condition in Experiment 1).

Participants

Overall 121 workers participated in the experiment: 47% were women, and 95% were native English speakers. Their average age was 28, and ranged from 19 to 48.

Design and Procedure

Participants were asked to generate a new idea after seeing a product idea (*See an example*), or generate a new idea using a schema, after seeing the schema and the product idea it was based on (the experimental conditions). Participants were either presented with a bad schemas (*Use a bad schema & see its example*) or good schema (*Use a good schema & see its example*). These schemas were randomly selected from the schemas generated by the crowd in Experiment 2. There were nine bad schemas (schemas rated 3 or lower on the 7-point quality scale), and nine good schemas (schemas rated 5 or higher). Examples of good and bad schemas are shown in Table 5. The product ideas these schemas were based on were the ideas shown in Table 1, one of which randomly showed up in the experiment.

Rating product quality

We used the procedure from Experiment 2 in which two raters blind to the experimental condition rated each product idea. The inter-rater reliabilities were .91, .80 and .81 for practicality, usefulness, and novelty respectively.

Analysis and Results

The means and standard errors of the ideas are shown in Figure 3. We tested hypotheses using a regression analysis with *quality* as the dependent variable, conditions as the independent variable, and education level and language as control variables. Results showed that ideas in *Use a good schema & see its example condition* were significantly better than those in *See an example*, $b = .78$, $t(76) = 2.64$, $p < .01$, as well as better than those in *Use a bad schema & see its example*, $b = .65$, $t(80) = 2.21$, $p < .05$. On the other hand, quality in the *Use a bad schema & see its example* and *See an example* were not significantly different, $b = .13$, $t(83) = .48$, $p = .64$. These results suggest that participants were more likely to generate good ideas using good schemas than seeing examples or using bad schemas. Table 5 shows

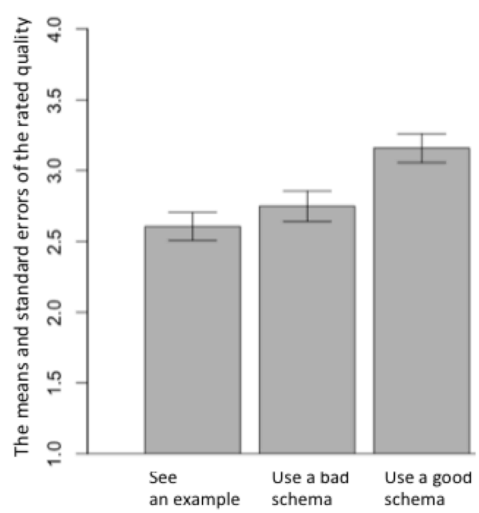

Figure 3. Experiment 3: the quality of the ideas

Good schemas	*Schema 1: The idea aims to attach from one unstable object to another stable object to keep the unstable object from moving.* Idea 1 (scored 4.40): Photographers can easily lose lens caps when shooting, especially if they're shooting in a crowded venue, such as a sporting event, or if they're shooting in nature. "Lens Cap Leaves" doubles as a lens hood and a lens cap. Imagine that the lens hood for a lens, rather than be one solid piece, is made up of interlocking leaves, like a vegetable steamer (here's an amazon link to what I'm talking about: http://www.amazon.com/Trudeau-Stainless-Steel-Vegetable-Steamer/dp/B00062B0K6). When you're done shooting, these leaves close up to seal the end of the lens, forming a cap that's permanently attached to the end of the lens. This will not only give the lens a permanently mounted lens hood, but will also give the lens a permanently mounted lens cap.
	Schema 2: Use an object to stabilize another object. It keeps it from moving around. Idea 2 (scored 4.56): My cat is very messy with her food and water. She is constantly knocking her bowls around, wasting food and water, creating a mess, and attracting ants and other critters. The bowls need to be held in place. A rubber base, specifically tailored to the size of common pet food bowls. The mechanism involves gripping tightly around the bowl, so as to not spill food and water in the base, as well as gripping to the floor, using a suction-cup system that leaves no trace, to ensure the bowls will be held in place.
Bad schemas	*Schema 3: The idea refers to something being secured to your head.* Idea 3 (scored 1.00): Too much light and sound during sleep time. Put on the eye and ear cover to get better sleep during the day.
	Schema 4: The idea aims to grab and hold. Idea 4 (scored 1.50): You are out protecting your hood with your gang of brothers and its late at night. You sit down for a quick rest, leaning up against the concrete wall. There is no way to get comfort in that position. What will you do? Now introducing the Power Nap Hoodie. You just raise your hood over your head and lean back. The extra fabric within creates a cushion between you and the wall.
Examples	Idea 5 (scored 2.00): Have you ever lost your ear buds or found them lying on a counter or table with a bunch of knots in them? This product attaches your waist strap or a belt loop of your pants. It is retractable ear buds. Never get knots or tangles in your ear buds again.

Table 5. Experiment 3: Examples of schemas and ideas. Scores range from 7 (best) to 1 (worst).

examples of ideas produced in different conditions.

In summary, Experiment 3 shows that participants using good schemas generated by others were more likely to generate good ideas than those who just saw examples. However, bad schemas did not have this effect. These findings further support the idea that distributed analogical idea generation is more effective than example-based idea generation. Importantly, together with Experiment 2, they demonstrate the viability of a crowdsourced, distributed analogy idea generation process, with one group generating schemas from examples and a second group generating ideas using the first group's schemas.

DISCUSSION

We introduced a novel approach for idea generation that leverages analogy to produce ideas better than those generated by just providing examples by spurring creativity in a structured way. We demonstrated the effectiveness of this approach through three experiments. Experiment 1 showed that people generate better product ideas when given good analogical schemas (in this case, experimenter-generated) than when given examples. Experiments 2 and 3 demonstrated that the crowd could generate useful schemas from examples, and that these crowd-generated schemas could be used by others to generate better product ideas than the examples that formed the basis of the schemas.

While the current work provides encouraging support for the value of using analogical transfer in creative ideation, it provides only one step towards understanding the cognitive processes involved in applying schemas in generating new ideas. Understanding exactly what makes a good vs. poor schema remains an important area for future research. For example, schemas containing both a purpose and a mechanism might provide people with "hammers" for searching for problems in different domains and therefore facilitate the searching process. If this is true, a good schema must have a good abstract purpose and a solution mechanism to assist this searching and mapping process. A bad schema, which doesn't meet such requirements (e.g., being too abstract without purpose or mechanism or being too concrete with many details of the examples), might fail to facilitate such a process. For example, in Table 5, Idea 3 and Idea 4 were generated using a very concrete schema and a very abstract schema, respectively. Idea 3 is overly similar to the given example and Idea 4 is not relevant to Schema 4 at all.

Why were schemas more effective than examples? Previous findings in analogy transfer model schema induction as a one-to-one mapping process of abstracting correspondences and eliminating differences [12]. This is consistent with our hypothesis that schemas abstract out surface features and thus might avoid fixation and conformity effects. For example, the "retractable ear buds" idea generated by a participant (Idea 5 in Table 5) is highly similar to one of the examples we provided, suggesting fixation at work. However, unlike studies on analogical transfer, the participants in our experiment did not have an analogous target problem to solve. Instead, they needed to search for a target domain, identify a candidate problem in that domain, and transfer the original schema to that problem in order to generate an idea. Given the challenges that previous research in analogical transfer have identified in promoting transfer even when the source and target problems have been hand-selected and perfectly alignable, we find it somewhat remarkable that crowds of untrained novices can accomplish this process consistently and show robust benefits.

One explanation for our success is that we are using realistic problems our participants could easily understand. Prior research has shown that people can accomplish complex abstract reasoning in familiar domains they would otherwise be unable to; for example, the notoriously challenging Wason selection task (testing if people understand the rule "if p then q") can be correctly solved if framed in a familiar domain (e.g., "if one is to drink alcohol, one must be over eighteen") [4]. If true, this would suggest that the crowd used for a given task should be familiar enough with the domain that they can induce the correct schemas for it. However, an interesting question is whether a small expert crowd (or the requester) might be able to generate a good schema and then use a larger, less expert crowd to generate ideas using the schema. Such an approach would maximize for expertise for inducing good schemas and maximize for diversity and scale for idea generation.

One question that merits further investigation is how to select appropriate examples (Step 1). In the present work we manually selected analogous and non-analogous examples in order to experimentally test the value of the approach. Promisingly, we did find it easy to find multiple analogous examples even among the relatively small selection of Quirky's existing products. However, scaling this process to the thousands of ideas that have been submitted to Quirky, let alone the millions of products in the U.S. patent database, is a formidable prospect. Some encouraging research shows that people are able to find examples based on deep analogical similarity even in Quirky's large product submission database [32]. Other crowdsourcing studies have shown the power of human workflows for classifying large datasets on deep structural features [e.g., 5, 1]. Such workflow designs, especially if combined with computational methods [29], might be a promising way to find ideas sharing common schemas. However, given our results that even given one example some participants could still generate good schemas, we believe our results can be useful even today for improving crowd idea generation.

Further Distributing Idea Generation

One significant advantage of schema-mediated idea generation is that it can be decomposed into subtasks, which can be further assigned to different individuals. As shown in our experiments, it is possible to implement the process with different crowds: one set of people for schema generation and another for idea generation. This process may be able to be broken down into even more sub-steps. For example, generating an idea from a schema involves multiple substeps which we combined into a single task: generating a target domain in which the schema may be applicable (e.g., "what are some domains in which objects need to be stabilized"), mapping the schema to that domain (e.g., "what are some possible objects in the new domain that map to elements of the *stability* schema"), and generating a solution (e.g., "this configuration of objects solves the problem in this domain using the *stability* schema"). One might implement these steps with different crowds and thereby benefit from their diversity.

Decomposability also allows for greater direction and management of the creative process. Each stage produces an output that is input to the next stage, meaning that the process can be restarted at any point to gather more options (e.g., schemas, domains, ideas). Quality control could be built into each stage, culling poor quality ideas before they propagate. Alternatively, if a requester is interested in a particular direction for ideas, schemas and domains that are of less interest could be pruned in order to focus attention on areas of greater relevance.

System design

These results from these experiments could inform the design of systems for crowd innovation. For example, an easy way to implement such an idea could be through a two-stage contest in crowdsourced innovation websites such as Quirky in which the crowd is asked to generate schemas and then the best schemas are chosen as the basis to generate ideas. Long term, one promising advantage of abstracting schemas from examples is that schemas could be aggregated and searched over. For example, the Props, Tether, and Eclipse examples in Figure 1 all share the *stability* schema. If such schemas were stored along with the examples then inventors could search for ideas based on schemas rather than being limited to keywords or surface attributes, finding potential solutions that could be transferred to their target domain. There are already many ideas on the Internet that could comprise the basis for such a schema-mediated search engine: Quirky alone has over 300,000 contributed ideas. Such a system could also increase the value of the many ideas that are not selected in design or ideation contests, making them potentially useful for sparking other ideas. Understanding how to represent schemas at the right level of abstraction and how to search for them remains an important issue for future research.

Limitations and Future Work

As with any initial study, there are many interesting questions left for future work. The participants in all our experiments were recruited from Mechanical Turk workers who are used to performing simple micro-tasks. One area for profitable future research is investigating whether this process will still be effective for experienced inventors or even teams of experts. On the one hand, it might be even more effective, because real inventors or experts could generate better schemas, domains, and ideas. On the other hand, they might be already using schemas in their process, limiting the additional gains.

Towards a Future of Improved Crowd Innovation

Today, most crowd inventors have little knowledge of, and few strategies for, effective innovation. Current crowd innovation sites typically reward only a tiny fraction of those contributing ideas, with the vast majority's contributions providing no benefits. With the growing number of publicly available idea repositories, it will be increasingly important and valuable to make existing ideas more useful for spurring new ideas. Such efforts might also help to address incentive issues: instead of only one idea which wins a contest being rewarded, imagine if unselected ideas that spur future ideas could be rewarded based on the success of the spurred ideas. Although more research is needed on how to abstract, represent, search across and integrate schemas to make such a future possible, the present research provides one step towards harnessing the power of crowd through analogy. We envision a future of interconnected crowd idea generation in which ideas and schemas build on and inspire each other, leading to a cycle of continuously accelerating innovation.

ACKNOWLEDGMENTS

This work was supported by NSF grants IIS-1149797, IIS-1217559, OCI-0943148, IIS-0968484, IIS-1111124, Bosch, Google, and Microsoft, and Carnegie Mellon's Center for the Future of Work.

REFERENCES

1. Andre P., Kittur, A., and Dow, S. P. 2014. Crowd synthesis: Extracting categories and clusters from complex data. In *Proc. CSCW*, ACM.
2. Alexander, C., Ishikawa, S., and Silverstein, M. A., *Pattern Language: Towns, Buildings, Construction.* Oxford University Press, 1977.
3. Buxton, B. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann, 2010.
4. Cheng, P.W. and Holyoak, K.J. Pragmatic reasoning schemas. *Cognitive psychology* 17, 4 (1985), 391–416.
5. Chilton, L., Little, G., Edge, D., Weld, D.S., and Landay, J.A. Cascade: Crowdsourcing Taxonomy Creation. In *Proc. CHI* 2013.
6. Clement, C.A. Effect of structural embedding on analogical transfer: Manifest versus latent analogs. *The American journal of psychology*, (1994), 1–38.
7. Finke, R.A., Ward, T.B., and Smith, S.M. *Creative cognition: Theory, research, and applications*. MIT press Cambridge, MA, 1992.
8. Fisher, R., *Statistical methods for research workers.* Oliver and Boyd, 1954.
9. Galton, F. The geometric mean, in vital and social statistics. In *Proc RSL 29*, (1879), 365–367.
10. Gentner, D. Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7, (1983), 155–170.
11. Gentner, D., S. Brem, R. Ferguson, P. Wolff, A. B. Markaman, K. and Forbus. *Analogy and creativity in the works of Johannes Kepler.* T. N. Ward, S. M. Smith, J. Vaid, eds. Creative Thought. Amer. Psych. Association, Washington, DC, (1997). 403–460.
12. Gick, M.L. and Holyoak, K.J. Schema induction and analogical transfer. *Cognitive psychology 15*, (1983), 1–38.
13. Gick, M.L. and Paterson, K. Do contrasting examples facilitate schema acquisition and analogical transfer? *Canadian Journal of Psychology* 46, 4 (1992), 539.
14. Hargadon, A.B. Brokering knowledge: Linking learning and innovation. *Research in Organizational behavior* 24, (2002), 41–85.
15. Hesse, Mary B., *The Structure of Scientific Inference*. Macmillan, London, 1974.
16. Hutter, K., Hautz, J., Füller, J., Mueller, J., and Matzler, K. The Tension between Competition and Collaboration in Community-Based Design Contests. *Creativity and Innovation Management* 20, 1 (2011), 3–21.
17. Hummel, J. E. and Holyoak K. J., Distributed representations of structure: a theory of analogical access and mapping. *Psychological Review* 104, (1997), 427–66
18. Hummel, J. E. and Holyoak, K. J., A symbolic-connectionist theory of relational inference and generalization. *Psychological Review* 110, (2003), 220–264.
19. Kittur, A., Chi, E.H., and Suh, B. Crowdsourcing user studies with Mechanical Turk. In *Proc. CHI'*2008, 453–456.
20. Kittur, A., Holyoak, K. J. and Hummel, J. E. Using idea observers in higher-order human category learning. In *Proc. CogSci*, *2006*.
21. Kittur, A.,Kraut, R. E. 2008. Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination. CSCW 2008. ACM.
22. Koh, T.K. *Essays on Technology-Enabled Platforms*. PhD Dissertation, Carnegie Mellon University, 2012.
23. Kulkarni, C., Dow, S.P., and Klemmer, S.R. Early and Repeated Exposure to Examples Improves Creative Work. In *Proc.* CogSci, 2012.
24. Lakhani, K.R. and Panetta, J.A. The principles of distributed innovation. *Innovations* 2, 3 (2007), 97–112.
25. Linsey, J., Laux, J., Clauss, E., Wood, K.L., and Markman, A. Increasing innovation: A trilogy of experiments towards a design-by-analogy method. Proceedings of the ASME 2007.
26. Reeves, L. and Weisberg, R.W. The role of content and abstract information in analogical transfer. *Psychological Bulletin 115*, 3 (1994), 381.
27. Smith, S.M., Ward, T.B., and Schumacher, J.S. Constraining effects of examples in a creative generation task. *Memory & Cognition* 21, 6 (1993), 837–845.
28. Sternberg, R.J. *Handbook of creativity*. Cambridge University Press, 1998.
29. Talton, J., Yang, L., Kumar, R., Lim, M., Goodman, N., and Mech, R. Learning design patterns with Bayesian grammar induction. In *Proc. UIST*, 2012, 63–74
30. Wolfgang, P., *Design patterns for object-oriented software development*. Addison-Wesley Publishing, England, 1995.
31. Yang, Y., Chen, P. Y., and Pavlou, P. Open innovation: An empirical study of online contests. In *Proc. ICIS* 2009.
32. Yu, L., Kittur, A., and Kraut, R. Searching for analogical ideas with crowds. *CHI'14*.
33. Yu, L. and Nickerson, J.V. Cooks or cobblers? Crowd creativity through combination. In *Proc. CHI 2011*, 1393–1402.
34. Yu, L. and Sakamoto, Y. Feature selection in crowd creativity. In *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems*. Springer, 2011, 383–392.