Measuring the Effects of Explicit Instructions and Incentives on the Idea Generation Rate of a Crowd-Based Population

by

Allan R. Donnelly

B.A. Architecture
University of California, Berkeley, 2006

SUBMITTED TO THE ENIGINEERING SYSTEMS DIVISION IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN ENGINEERING AND MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2013

Signature redacted

Signature of Author: ____________________
Engineering Systems Division
23 May 2013

Signature redacted

Certified by: ____________________
Thomas W. Malone
Patrick J. McGovern Professor of Management
Thesis Supervisor

Signature redacted

Accepted by: ____________________
Pat Hale
Engineering System Division
Program Director, System Design and Management

Measuring the Effects of Explicit Instructions and Incentives on the Idea Generation Rate of a Crowd-Based Population

by

Allan R. Donnelly

Submitted to the Engineering Systems Division
On 23 May 2013 in Partial Fulfillment of the
Requirements for the Degree of Master in Science in
Engineering and Management

Abstract

Management researchers have long sought strategies for increasing the rate and quality of ideas generated among workers. Additionally, the advent of internet-based communications has created opportunities for valuable ideas to be generated – and harnessed – from crowds of individuals. The first section of this paper reviews the early and recent literature on measuring creativity, focusing specifically on the effects of incentives and explicit instructions on the rate of idea generation. The following section describes the crowdsourcing platform through which the research was conducted – Amazon Mechanical Turk (AMT) – and reviews recent work that has utilized this platform for experimental research. The project at hand engages participants in a divergent thinking exercise to measure the rate of idea generation for the crowd-based population. The findings show more unique ideas occur later in the response period, demonstrating the presence of the serial order effect; that explicitly instructing respondents to "Be Creative" increases the rate of idea generation; and that offering a bonus incentive for "especially creative ideas" decreases the rate of idea generation for specific demographics of respondents. The paper continues with a discussion of research limitations and areas for further exploration. Conclusions and insights are offered at the end.

Thesis Supervisor: Thomas W. Malone
Title: Patrick J. McGovern Professor of Management

This Page Intentionally Left Blank

Table of Contents

Literature Review

Measuring Creativity

Generating new ideas is at the core of innovation. As the industrial economies of the mid-twentieth century gave way to large and complex organizations, management researchers began exploring methodologies for measuring the rate at which creative – and innovative – ideas are generated (Guilford, 1950; Guilford, 1951; Wilson et al., 1954; Christensen et al., 1957). This work sought to link certain tests with specific measures of idea evaluation. Eight measures of idea evaluation were defined in this work: *Sensitivity to Problems, Fluency, Flexibility, Originality, Penetration, Analysis, Synthesis*, and *Redefinition*. Each of these metrics aimed to assign some value to the rate at which respondents generated ideas or the quality of the ideas generated. Values were then combined and synthesized to evaluate the psychological profiles of individual respondents. Guilford, Wilson, and Christensen's work laid the groundwork for future research pursuing to measure and evaluate creativity and creative ideas.

One of the studies used by these researchers to test the generation rate of creative ideas was the unusual uses test or the brick uses test. This test asked respondents to provide as many unusual uses for a brick as they could think of in the time allotted. Asking an open ended question of this nature allowed researchers to chart the rate at which research subjects generated responses by simply counting the overall number of ideas generated. This metric is called *fluency* and has been measured by many researches since (Harrington, 1975; Ward et al., 2004). Christensen at al. (1957) found respondents who had higher fluency rates for the unusual uses test had greater *Flexibility* in their thinking. *Flexibility* in this work is defined as the cognitive ability to move between different domains of responses. A respondent might start off by listing all the uses for a brick that might involve building. And then, once he has exhausted ideas within that domain, he might switch to uses for a brick that involve breaking other objects. Respondents who were able to switch between multiple domains were also able to come up with a greater number and quality of unique ideas.

The Serial Order Effect

An important contribution of the early work of Christensen, Guildford, and Wilson is that it is the first to document the serial order effect (Christensen et al., 1957). The serial order effect refers to the fact that as the test period progresses, ideas generated tend to become more original, novel, and remote even as the fluency of ideas generated decreases (Beaty and Silvia, 2012). Christensen et al. (1957) also used the unusual uses test to measure this effect. The work showed that as time passed, fluency decreased and novelty increased. Many researchers have since measured and demonstrated the serial order effect, charting its existence across various populations (Christensen et al. 1957, Chen et al., 2005; Beaty and Silvia, 2012). Beaty and Silvia (2012) confirm the existence of the serial order effect within their research population of young adults, and find that intelligence can diminish the serial order effect (i.e. respondents with higher intelligence scores generated unique ideas throughout the response period).

As management science progressed, researchers started placing more emphasis on the evaluation of the relative creativity of an idea rather than just the number of ideas generated. The early research is also important here as the factors defined by Christensen, Guilford, and Wilson became the basis for the formulation of future cognitive and intelligence tests (Vernan, 1960; Ekstrom, 1979), employee evaluation surveys (Ruch, 1980), and idea evaluation frameworks (Amabile, 1983; MacCrimmon and Wagner, 1994; Dean et al., 2006). Recent idea evaluation frameworks have evolved these single measures of creativity into multi-attribute definitions of creative products. MacCrimmon and Wagner (1994) create a basis for defining creative ideas as those that are both *Novel* and *Quality* (MacCrimmon and Wagner, 1994). *Novel* ideas are defined as rare, unusual, or uncommon. *Novelty* is the evolution of Guilford's metric of *Originality*. In fact, MacCrimmon and Wagner (1994) use *Originality* coupled with *Paradigm Relatedness* – an extension of Guilford's measure of *Redefinition* – as sub-attributes by which *Novelty* is measured, further illustrating the evolution of Guilford's work in current research.

One challenge with the current body of research on creative idea evaluation is that different researchers define and measure dimensions of creativity in different ways. Dean et al. (2006) provides an exhaustive review of 91 of the most recent articles studying creative idea evaluation. Upon assessing the historical research, Dean et al. distill commonly used constructs, create reliable scales for each construct, and provide an approach for aggregating measures of creativity. This work incorporates several historical constructs into one comprehensive measure of creativity and provides a methodology for future researchers to follow when evaluating creative ideas.

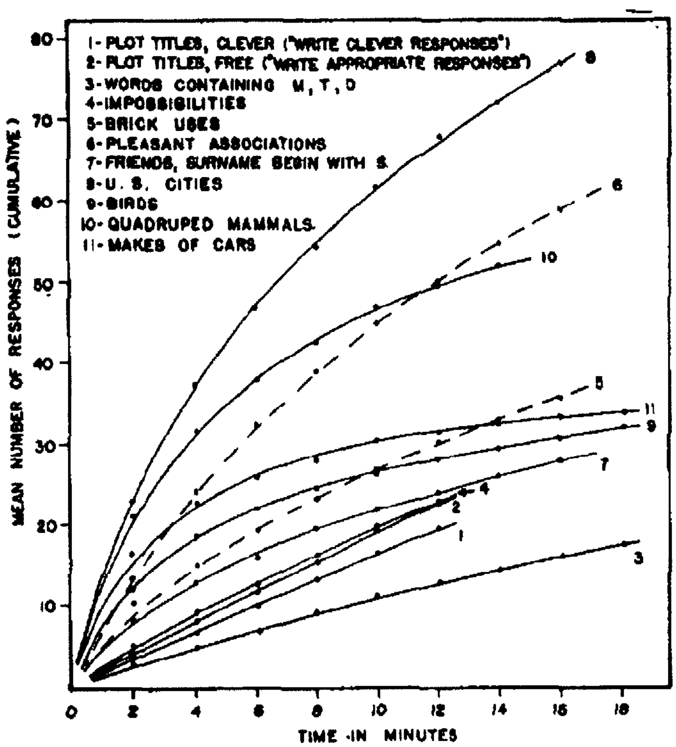
The process of evaluating creative ideas can be somewhat subjective, so this framework and methodology creates an objective measure for defining creativity. In short, Dean et al. define creative ideas as those that are both *Novel* and *Quality*. *Novelty* is calculated by assigning values to two sub-dimensions: *Originality* and *Paradigm Relatedness*. *Quality* is deconstructed into three dimensions: *Workability*, *Relevance*, and *Specificity*. *Workability* is further deconstructed into sub-dimensions of *Acceptability* and *Implementability*; *Relevance* into *Applicability* and *Effectiveness*; and *Specificity* into *Implicational Explicitness*, *Completeness*, and *Clarity*. Dean et al. test their creativity evaluation measures in both online and manual experiments to demonstrate the use of their comprehensive framework and to propose an evaluation system to improve comparability of future studies.

While the majority of researchers seeking to evaluate creative ideas utilize some form of evaluative framework that is defined by either one holistic definition or a multi-attribute definition of creativity (Dean et al., 2006), other researchers measure the rate and quantity of ideas generated without seeking to evaluate the creativity of responses (Firestein, 1990; Dennis & Valacich, 1993; Crowne & Ross 1995). In this work, the primary metric is *Fluency*, or the total number of non-duplicate ideas generated (Guilford, 1954). Some researchers have found a correlation between fluency and novelty (Briggs et al, 1997) meaning that as the number of ideas generated increases, the quality of ideas generated also tends to increase. Although ideas generated later in the response period might be more *novel* than those generated earlier in the response period, later responses are not necessarily of higher quality than earlier responses. This has lead recent researchers to test and identify strategies that increase both the quantity and quality of ideas generated.

Effects of Question Framing on Creative Idea Generation

Another major finding in the early research on creativity suggests that using instructions that explicitly ask respondents to think creatively has the potential to increase quality, but not necessarily the fluency, of ideas generated (Exhibit 1) (Christensen et al., 1957). Christensen el at. (1957) hypothesize that the decrease in overall idea production under specific instruction to be creative might encourage the respondent to self-sensor ideas that he thinks are less than creative or might even hinder him from thinking of creative ideas to begin with. This has led contemporary researchers to explore techniques that might increase the rate and quality of ideas generated. Some have used various methodologies to affect the rate of idea generation including altering question framing (Harrington, 1975; Heylighen, 1988; Eisenberger and Armeli, 1997; Ward et al., 2004; Niu and Liu, 2009), providing creativity training (Firestien,1990; Scott et al., 2004; Osburn and Mumford, 2006), offering financial incentives (Amabile, 1983; Hennessey and Amabile, 1988; Eisenberger and Selbst, 1994; Eisenberger & Armeli, 1997; Eisenberger & Rhoades, 2001; Eisenberger & Shanock, 2003; Toubia, 2006; Borst, 2010), and forecasting implications and effects of ideas generated (Byrne et al., 2010). This research project aims to test the effects of altering question framing and offering financial incentives in order to change the rate at which ideas are generated.

Exhibit 1: Production Curves of Cumulative Number of Responses As a Function of Time (Christensen et al., 1957)



Researchers following the work of Christensen, Guildford, and Wilson further explored the effects of explicit instructions to be creative on the rate and quality of ideas generated. One such study found that using explicit instructions to be creative increased both the quantity and creativity of responses provided (Harrington, 1975). Contemporary research in this domain broadened the findings to show that explicit instructions to think more abstractly (rather than creatively) led to more novel creations (Ward, 2004). Further work was conducted to understand if placing an emphasis on creative thinking would demonstrate similar effects on the rate of ideas generated by respondents from a non-western population (Niu and Liu, 2009). This work found that simply prompting respondents to think creatively did not enhance their creative output, but that more elaborate instructions did increase the creativity of responses. This work demonstrates a clear connection between explicit instructions to think creatively or abstractly and an increase in the quality and quantity of ideas generated.

Effects of Monetary Incentives on Creative Idea Generation

Another study that found a link between explicit instructions to think creatively and the creative performance, also found that coupling explicit instructions with a large reward increased participants creative output and effort in subsequent and unrelated exercises. (Eisenberger and Armeli, 1997). The introduction of an incentive into the creative thinking process has complex effects on the motivations and performance of a person. As with many areas of management and social science, the research on creativity explores what the effects of monetary and other incentives and rewards might be on the quantity and quality of ideas generated. There is some debate as to whether the presence of a reward increases or decreases creative idea generation. Some researchers suggest that rewards can have a negative impact of the rate of ideas generated (Amabile, 1983, Hennessey and Amabile, 1988). Studied the effects of rewarding stuendts for one task on their creative performance on subsequent tasks. This work found that repeated reward of a given task would foster an expectation of future rewards and a reduction in the subject's creative performance on subsequent tasks (Hennessey and Amabile, 1988).

A review of the more current research on the effects of rewards on creative idea generation suggests that offering rewards can increase the quantity and quality of ideas generated (Eisenberger and Selbst, 1994; Eisenberger & Armeli, 1997; Eisenberger & Rhoades, 2001; Eisenberger & Shanock, 2003; Toubia, 2006; Borst, 2010). The various works authored by Eisenberger focus mainly on the effects of rewards and incentives on the creative performance of preadolescent school children. In the first of these studies, it is shown that for divergent thinking tasks, smaller, rather than larger, rewards can increase creativity (Eisenberger and Selbst, 1994). A subsequent study demonstrated that repeatedly rewarding students for high levels of creativity had the effect on increasing their creative performance on certain subsequent tasks (Eisenberger and Rhodes, 2001). An additional study supporting the notion that incentives can increase creative performance involves the use of a "web-based, asynchronous ideation game" and finds that the presence of an incentive motivated participants to try harder to generate ideas and to give up less easily (Toubia, 2006). These works demonstrate that certain incentives and motivations have the capacity to effect the intrinsic motivations of school children.

While these studies come to similar conclusions about the ability of incentives to increase creative performance, the various methodologies and differences in populations tested make it difficult to compare results across the research findings. Some researchers, as described above, tested the impact of incentives by studying the creative output of tasks subsequent to the offering of a reward (Eisenberger et al., 1994; Eisenberger & Armeli, 1998; Eisenberger & Rhoades 2001; Eisenberger & Shanock, 2003). Other researchers engaged subjects in nominal groups – in-person groups where discussion amongst group members is encouraged (Diehl and Stroebe, 1991; Peeters et al., 2010) – or group brainstorming activities (Linsey et al. 2005). Whereas others still, tested the effects of rewards on idea generation using online communication tools (Borst, 2010; MacCrimmon and Wagner, 1994). This work does suggest that incentives have the potential to increase creativity, but further work is necessary to understand what, if any, causal relationship exists between incentivizing creativity and idea generation.

Computer Collection Systems

Another important aspect of the current research on creative idea generation explores how the medium through which subjects are engaged and ideas are collected affects the rate and quality of ideas generated. Early creativity research directly engaged individuals as subjects in the generation of ideas (Guildford, 1950; Wilson et al., 1957). Other research has explored the effects of group decision-making and group ideating behaviors on the rate and quality of ideas generated (Diehl and Stroebe, 1991; Linsey, 2005). One problem identified in the research on group ideation points to an apparent productivity loss when group performance is compared to that of a nominal group – sets of individuals working alone (Diehl and Stroebe, 1991; Paulus, 2000; Nijstad *et al, 2003;*). Some research has shown that nominal groups generate both higher quantities and quality of ideas (Dennis and Valacich, 1993; Dennis and Valacich, 1994). This work was conducted using in-person collection methodologies. Diehl and Stroebe conducted a series of tests to identify the mechanisms that mediate the impact of production blocking on the productivity of idea-generating groups (Diehl and Stroebe, 1991). Their research found that increasing the amount of time allotted for ideation had the directly proportional effect of increasing the number of ideas generated. It also found that speaking time did not have an effect on the number of ideas generated.

As computer-based and internet-enabled communication systems became available, researchers started testing how the use of computers to engage individuals, groups, and nominal groups in the creative idea generation process might affect the rate and quality of ideas generated (Dennis and Valacich, 1993; Dennis and Valacich, 1994; MacCrimmon and Wagner, 1994; et al. s, 1997). This research suggests that using computers as a medium for idea collection and group interaction can mitigate the productivity loss found in in-person group settings (Dennis and Valacich, 1993; MacCrimmon and Wagner, 1994). Dennis and Valacich find that groups interacting through a computer mediated electornic brainstomring medium generated more ideas than nominal groups of the same size (Dennis and Valacich, 1993). In a subsequent sutdy, these researchers also found an increase in both the quantity and quality of ideas generated by computer-mediated groups (Dennis and Valacich, 1994). Another study examining the effects of computer-mediated brainstorming on the quality of ideas generated found similar increases in the creativity of ideas generated (MacCrimmon and Wagner, 1994). This study also found an amplification of creativite idea generation in the subjects defined as highly creative. Together, these studies suggest that the use of computers as communication tools might lessen the impacts of process losses such as production blocking, evaluation apprehension, and free riding (Dennis and Valacich, 1993).

Recent Crowdsourcing Research

As global access to online communications has grown, new kinds of online populations have emerged. Masses of individual computer users dispersed throughout the world join together into crowd-based communities over the Internet. These crowds are a relatively modern manifestation of collective intelligence. Collective intelligence refers to the ability of a group to solve more problems than its individual members (Heylighen, 1999). It is suggested that collective intelligence is aided by the fact that computer based systems have the ability to decrease process loss effects on very large groups of people (Heylighen, 1999). To this end, computer based platforms like, Amazon Mechanical Turk, OpenIDEO, Innocentive, and Top Coder have been built to host and harness the collective intelligence of crowds of workers.

The population of respondents for this research came from Amazon's Mechanical Turk (AMT) marketplace for human intelligence work. AMT is a crowd-sourced marketplace that connects people requesting human intelligence tasks (HITs)-to-be-done with a network of hundreds of thousands of human intelligence workers. Requesters seek information that cannot be generated by a computer and must therefore be created by a human. Requesters create HITs that require workers to fill out surveys, find and enter data, review images, and transcribe audio/visual recordings. For completing a HIT, requesters also offer a cash reward ranging from $0.01 for very quick tasks and up to $25 for transcription tasks that could take an hour or more.

Although AMT was designed as a crowd-sourcing platform for knowledge-based tasks, some researchers have started utilizing the network as a source of subjects for various projects (Erikson & Simpson, 2010; Mason & Watts, 2009; Suri & Watts, 2011). There are three major benefits to using AMT over other Internet enabled research collection methodologies: Subject pool access, subject pool diversity, and low cost (Mason & Suri, 2012). AMT platform allows researchers an easy medium to conduct longitudinal research. AMT provides a robust user platform that is flexible enough to enable researchers across a wide range of domains the opportunity to engage participants in important and valuable research. Moreover, as a crowd-based platform, AMT creates cheap and easy connection between individual research teams and hundreds of thousands of potential respondents.

Many of the studies conducted using AMT workers as a population of respondents have little relationship to the topics explored in the research at hand. This is likely due to the fact that AMT and other crowdsourcing platforms have only recently begun to be used for social science research. One study that does have some relevance to the present research explores the relationship between financial compensation and worker performance (Mason & Watts, 2009). In this work, Mason and Watts find that larger financial rewards increased the quantity but not the quality of work. Furthermore, their work shows that a quota system for awarding payment results in better performance for less pay than an equivalent piece rate system. (Mason and Watts, 2009) While the tasks and activities used to facilitate this research exercise differ from the divergent thinking tasks of the present study, the findings of Mason and Watts suggest that large rewards for creative thinking may decrease the quality of responses provided.

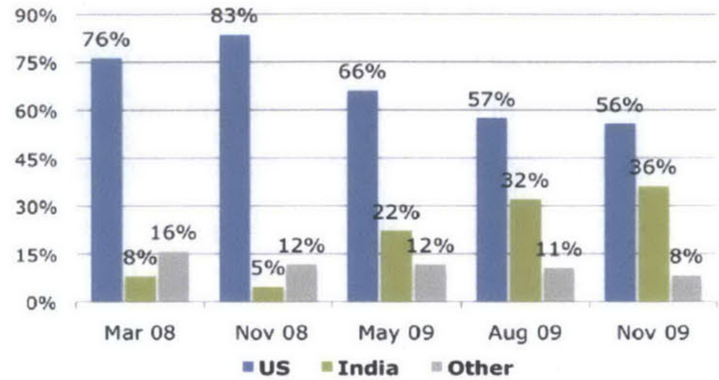
Demographics of Mechanical Turk Population

The most recent survey testing the demographics of the AMT population was published in 2010 (Ross et al., 2010). This research was conducted longitudinally over the course of one year beginning in November 2008 and ending in November 2009. The study found several trends that indicated rapid changes in the the AMT worker population. The major finding from this work is that AMT workers tend to be younger and more highly educated than the general public. A majority of workers fall under the age of 35 and hold a bachelors or graduate degree (Ross et al., 2010). The study also found shifts in the gender, and nationality distributions of AMT workers. Over the course of one year, the percentage of AMT workers in India grew from 8% to 36% while the percentage of workers in the United States fell from a high of 83% to a low of 56% (Ross et al., 2010). Additionally, the survey data indicates a growing proportion of male workers in the AMT population. This study finds that between in 2009, the population of AMT workers was shifting from a workforce consisting primarily of moderate-income, US-based workers to one consisting of young, highly-educated workers from India (Ross et al., 2010).

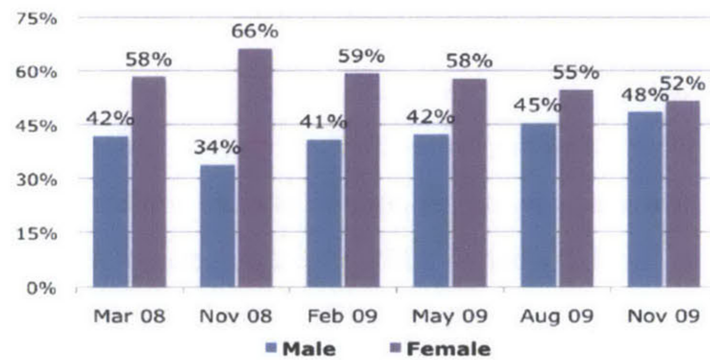
It is important to note that the demographics of the AMT population described in this study are out of date and are not precisely descriptive of the population of AMT workers at the time of this research. More up-to-date research on the demographics of the AMT population is needed in order to understand how the population of respondents to this research might be representative of the population of AMT workers as a whole.

Exhibit 2: Demographic Data of AMT Worker Population Nov 2008 – Nov 2009 (Ross et al., 2010)

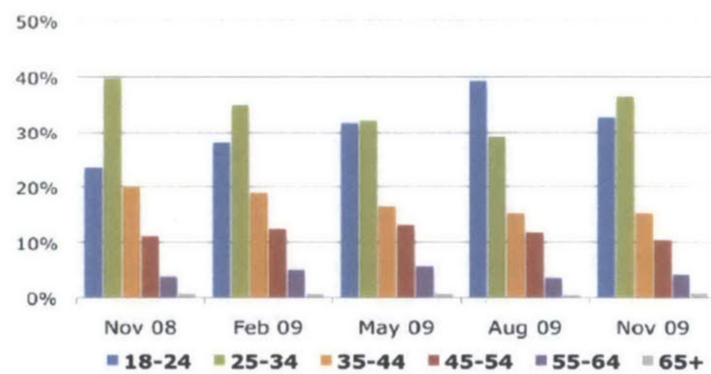
Nationality of MTurk workers over time. Countries with more than 1% of respondents include Canada, the U.K., and the Philippines. We do not have country data for Feb. 2009.



Gender of MTurk workers over time.



Age of MTurk workers over time.



Methodology

Hypotheses

It is the aim of this research to identify strategies for motivating a group of workers to generate as many ideas as possible. To this end, the research measures the fluency of responses to a divergent thinking task along with the effects of explicit instructions to think creatively and the appearance of monetary incentives on the fluency of responses generated. Three main hypotheses were formulated for this portion of the analysis based on the content and findings of the literature on creative idea evaluation and generation.

Serial Order Effect

The first hypothesis is based on the work demonstrating the serial order effect. From this work it is expected that the quantity of ideas generated will decrease over time and that more novel responses will occur later in the response period.

Question Frames

Hypotheses two and three test the effects of altered question frames on the fluency of new responses. The second hypothesis expects that explicit instructions to "BE CREATIVE" will increase overall fluency, while the third hypothesis expects the appearance of a monetary incentive will also increase the fluency of responses.

Demographic Sorting

In addition to identifying strategies for increasing the idea generation rate of a group of workers, this research also analyzes the relationship between demographic factors and fluency. This part of the analysis simulates various nominal-group compositions based on the demographic factors provided by research participants. Two hypotheses are stated for this section. First, workers based in the United States are expected to generate more new ideas than those based in India. Second, it is hypothesized that female respondents will provide more new ideas than male respondents. Finally, based on various findings of the analysis, suggestions for ideal group composition are made.

Responses and Respondent Population

Research subjects were engaged using the AMT human intelligence platform. The assignment was titled "Answer a Simple Survey and Provide Responses to a Simple Question." Respondents were asked to fill out a short survey providing demographic information regarding their age, gender, nationality, level of education, and area of expertise. Respondents were then engaged in a variation of the unusual uses test where they were asked to provide ten responses to the question "What can you do with a brick?" The control question – "What can you do with a brick?" – was developed from the work of Guilford et al. 1950, 1954, 1957. An additional framing of this question was posited in order to test the effects of explicit instructions to be creative: "What can you do with a brick? (BE CREATIVE)." A third question frame was posited to test the effects of a potential monetary incentive: "What can you do with a brick? (Especially creative answers may receive a bonus)." Subjects filled in separate open-ended response boxes numbered one through ten. In return for completing a HIT (i.e. providing demographic information and ten responses) workers were paid $0.10.

A total of 325 HITs were completed across the three question frames – 113 for the control frame, 112 for the be-creative frame, and 100 for the incentive frame. HITs that were not fully completed (i.e. respondents provided fewer than ten answers to the given question) were discarded. Additionally, HITs tagged with the same Worker ID number were filtered by response submission time and only the first completed HIT was retained. Filtering the responses in this way yielded a total of 284 respondents to the three question frames: 95 control, 95 be creative, and 94 incentive. This provided a total of 2,840 individual responses to the general question "What can you do with a brick?"

Since the goal of this research is to test the AMT crowd as a population, a series of requirements were created in order to test a wide range of AMT workers. AMT workers are assigned a rating based on how many HITs they complete and how many of those HITs are accepted by the requestor. The AMT default settings allow only Master Workers (i.e. workers who demonstrate excellence in a certain HIT type) to complete the task. Because this research intended to test the fluency of the crowd population, a set of customized worker requirements was created. To be eligible to complete a HIT for this project, respondents had to have a HIT Approval Rate for All Requester's HITs of at least 95%. The Requester profile for this project was set up as new account, so all AMT workers had the opportunity to complete a HIT. Additionally, Workers had to have at least 100 approved HITs under their belt in order to complete this project. This is a relatively low threshold for AMT workers and it seeks only to exclude new and/or non-productive workers.

Respondent Demographics

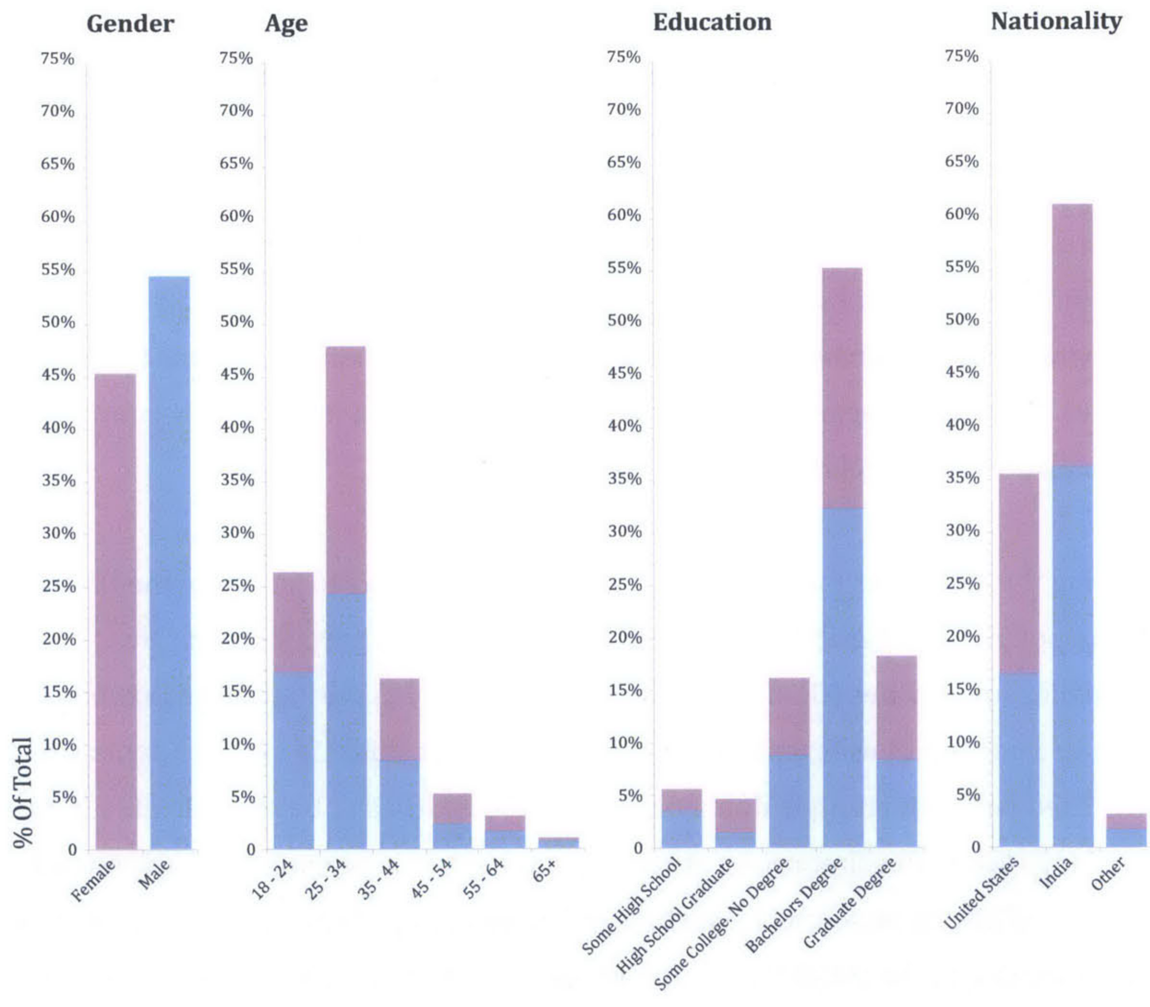
The demographics of the respondent population (Exhibit 3) differ from those charted in the most recent study of the AMT population (Ross et al., 2010). This difference is likely due to the fact that the most recent AMT demographic information comes from a longitudinal study conducted from 2008 to 2009, and the AMT platform and worker population have likely evolved since then. The demographics of the respondent population may be representative of the AMT population as a whole, but without more recent data, conclusions about representativeness cannot be made.

While the age rage of the respondent population spanned from 19 to 69, the response population shows a high percentage of AMT workers in the under the age of 35 (74.3%). While this skew toward younger workers is consistent with the earlier demographic research findings, the percentage of workers under 35 in the respondent population is much higher than the maximum percentage shown in any of the segments of the longitudinal study (e.g. 68% in November 2009) (Ross et al., 2010). The age demographics of this project more closely align with those found in another recent study using AMT worker (Mason and Suri, 2011). If the respondent population is characteristic of the AMT population as a whole, then this suggests that the population of AMT workers has grown much younger. The response population also shows more male respondents (54.6%) than female respondents (45.4%). The 2010 data does illustrate a shifting trend in the gender of AMT workers, showing a steady rise in the percentage of male workers. But as of November 2009, females still outnumbered males in the AMT worker population.

Another major difference between the respondent population and the 2010 demographic information is a heavy skew toward workers from India. The 2010 data shows a rise in the percentage of workers from India. But it also shows US workers as more than half the worker population (56%). The respondent population was primarily comprised of people from the United States and India with nearly twice as many respondents from India (61.3%) as from the United States (35.6%), and only 3.2% of respondents from a different country of origin. Again, it cannot be determined from this data whether or not the research population is representative of the population as a whole, but the data does suggest that the AMT population has change drastically since the most recent surveys.

One area where the demographics of the research population appear to be consistent with the 2010 data is in the level of education achieved by AMT workers. Both the 2010 data and the research data show a highly educated workforce. In the research population, the percentage of workers from the United States holding a Bachelors degree is higher than that of the AMT population, but the percentage of workers from the US holding graduate degrees is about the same. The same is true of the population of workers from India: Bachelors degrees have increased and graduate degrees have stayed about the same. Appendices 1 and 2 chart the distribution of gender and education level of the respondent population across the other demographic measures. This shows that highly educated, young Indian men are the primary respondents for this research.

Exhibit 3: Respondent Demographics



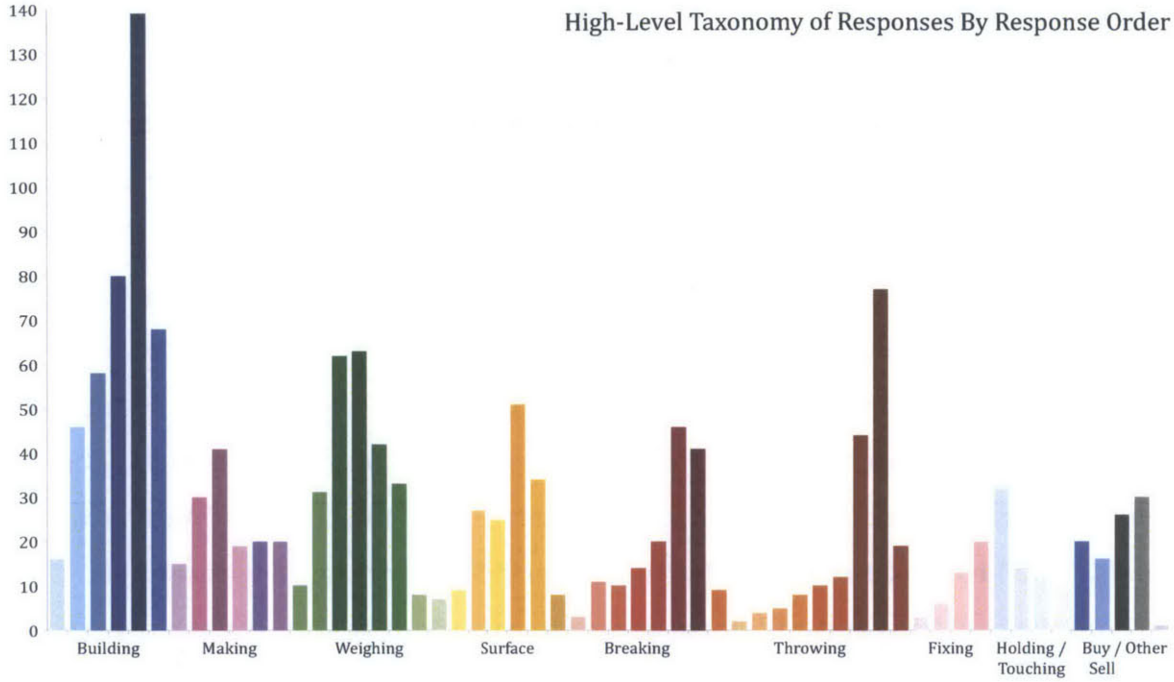
Response Tagging

In order to analyze the fluency of responses, each new response to the brick uses question had to be identified and tagged. Tagging began by sorting responses into categories of similar uses. For example, all of the responses pertaining to building and construction were clustered together into a high-level use category called *Building/Making*. This category of responses was then sorted into subcategories like *Build House* and *Build Non-House Structure.* Other such categories in this taxonomy of brick uses include *Throwing, Breaking, Sitting, Stepping, Weighing* etc. This high–level sorting yielded a list of 14 categories and 81 subcategories for the 2,840 responses (see Appendix A for a taxonomy of high level categories and subcategories). Exhibit 4 illustrates the total number of responses for each of the subcategories in the high-level taxonomy. Responses proposing using the brick to *Build a Non-House Structure* were the most common responses in this sorting. Exhibit 4 also illustrates the same high-level categorization with added detail showing the total number of responses in each of the response positions (e.g. 1, 2, 3…10). This chart suggests that responses that occur with less frequency also tend to show up later in the response period.

Sorting the responses into high-level categories is one way to illustrate the *flexibility* of the population. *Flexibility* is defined as the capacity to get out of a rut by switching approaches (Guilford, 1950). Each category in the high-level taxonomy is akin to a domain of uses that a respondent might "get stuck" in. Respondents who switched between multiple domains might be considered to be more flexible in their thinking. While this first round of sorting was useful for beginning to understand the rate of idea generation, the sorting and categorizing of responses was somewhat arbitrary due to the subjective nature of likeness. Additionally, responses in one category could be different enough to warrant the creation of a new category. Take, for example, two responses that were both grouped into the *Build Non-House Structure* category: *Build Mosque* and *Build Fort.* While these two responses were appropriate under the heading of *Build Non-House Structure*, they were different enough to suggest that another grouping might be more appropriate. For this reason, more granularity was needed in the response tagging process in order to accurately identify each new idea and measure the response generation rate.

To this end, several additional rounds of tagging and sorting were conducted, yielding a taxonomy of 29 categories and a list of 732 tags for the 2,840 responses. Tags were developed for each of the question frames separately and were then compiled into a master list of 732 tags. Basic responses were generally tagged using a two-word phrase describing the action and object of the response. For example the responses "Build a house'" and "Use the brick to build a house" were both tagged using the action-object pair *Build House* (see Appendix B for a complete list of tags and responses). *Build House* is, however, separate from the tags *Build Mosque* or *Build Fort.* Responses that included more specificity were tagged with a basic action-object pairing and additional descriptive words. For example, the response "Make artistic picture on brick" was tagged with the action-object pair of *Art On [Brick]* whereas the response "Draw art on the bricks and sell them on ebay" was tagged with the same action-object pair and an additional modifier *Art On Brick Sell On Ebay.* The 732 tags represent an exhaustive list of each new response generated across the three question frames (Exhibit 4).

Exhibit 4: High Level Taxonomy of Brick Uses Responses By Category and By Response Order



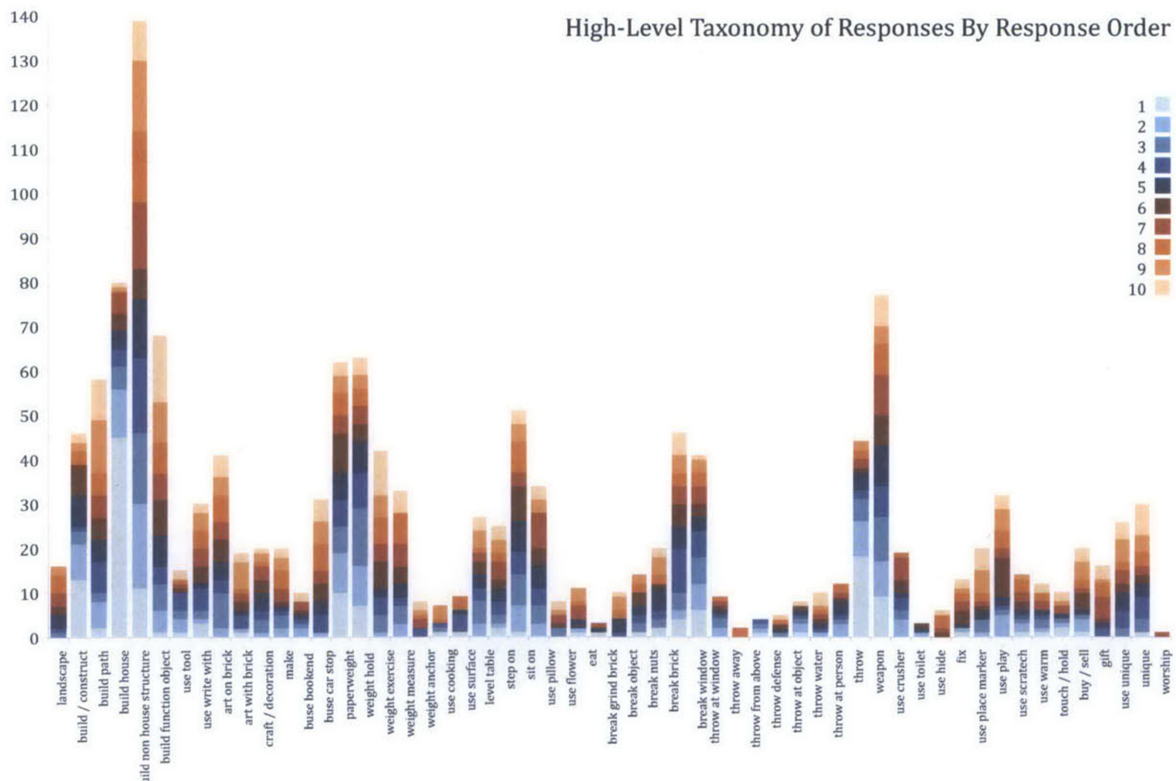
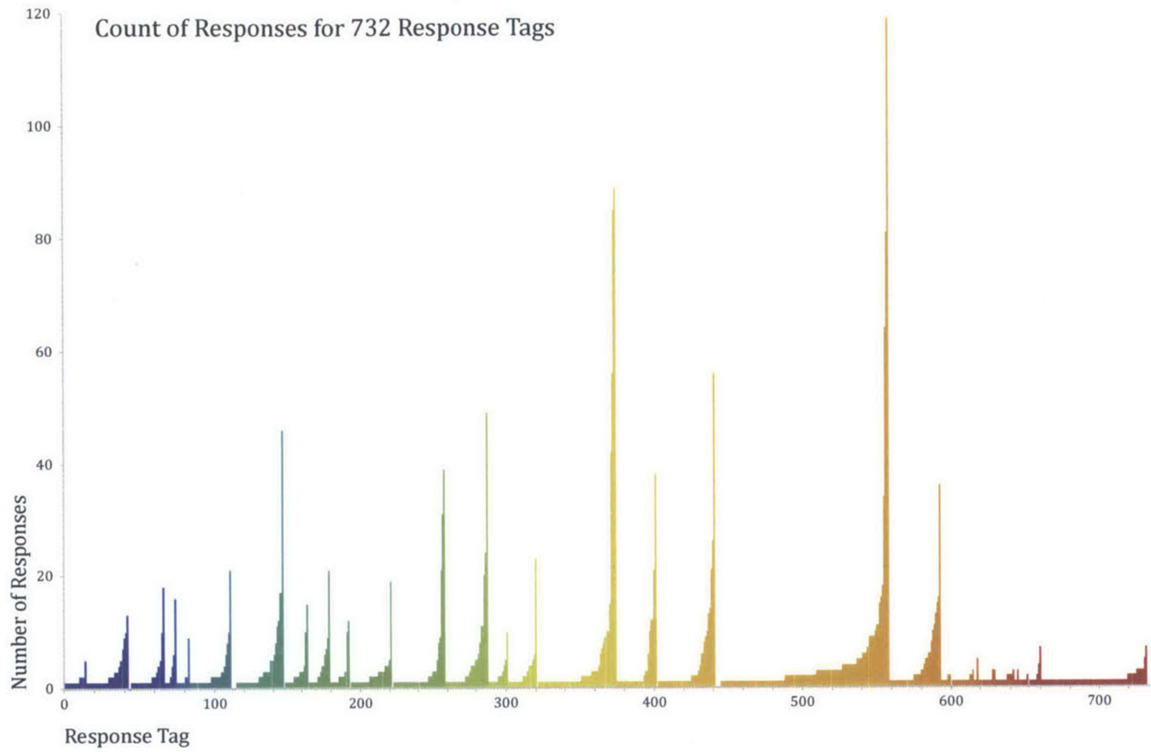


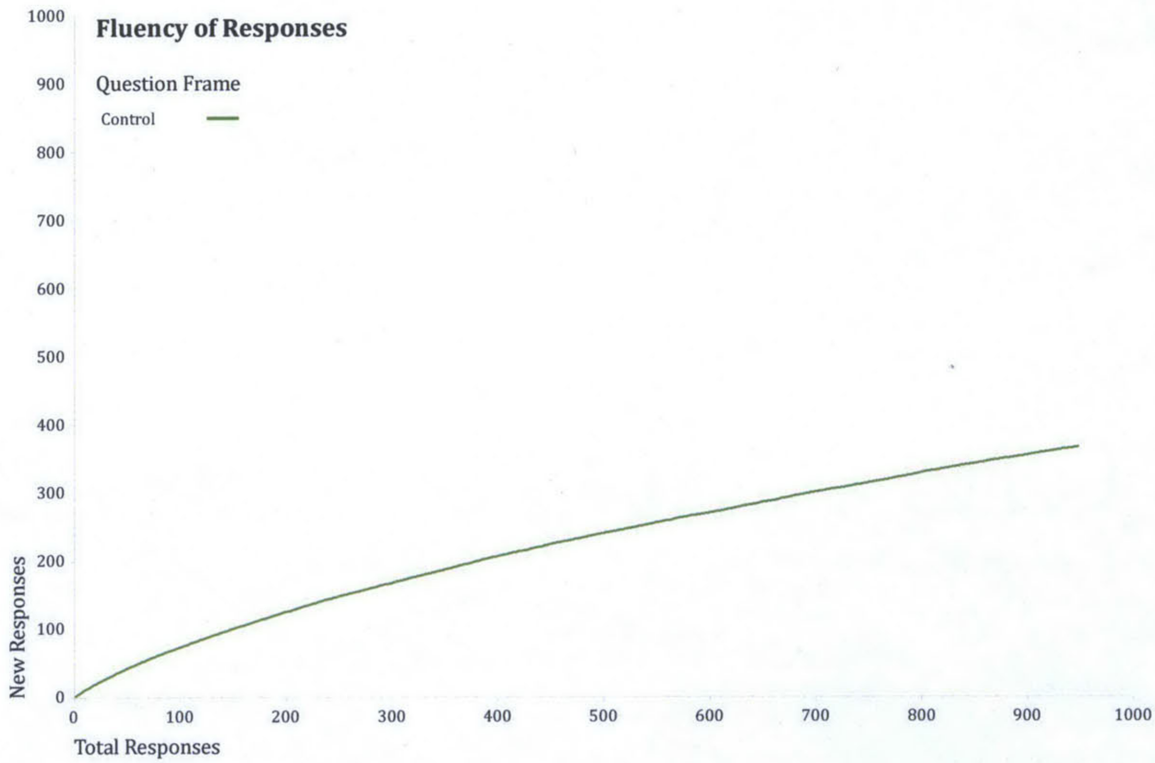
Exhibit 5: Count of Responses for 732 Response Tags



Simulation

To chart the fluency of responses from AMT workers, a cumulative summation of all new responses had to be calculated. An exact curve illustrating this summation is difficult to chart because the order in which responses are counted is arbitrary. That is to say, the discrete intervals between new ideas can vary depending on how the data is sorted. For instance, sorting respondents by age in ascending order will illustrate a different curve than sorting respondents by age in descending order. This is due to the variability in the distribution of responses per response tag. In order to preform statistical analysis therefore, it was necessary to simulate many possible response sequencings and then test the sets of numbers against one another for to assess significance. To run the simulations, a constantly updating random number was assigned to each of the responses. The responses were then sorted using this random number (and other parameters), and the cumulative fluency of the simulation was calculated. Thirty simulated sample sets were generated and an average summation was calculated (Exhibit 6).

Exhibit 6: Fluency of Responses For Control Question Frame



The arbitrary nature of the response sorting also provides the opportunity to compose hypothetical groupings of responses based on respondent demographics. Simulating the idea generation rate of possible groupings of respondents might suggest certain relationships between group composition and idea generation. Furthermore, the random ordering of responses assisted in the demonstration of the serial order effect. To this end, thirty simulations were generated for each of the following data sets in order to test the aforementioned hypotheses:

Control Responses
Explicit Instructions Responses
Monetary Incentive Responses
Serial Order Responses
Serial Order and Education Level Responses
Responses By Gender

A t-test was used to calculate the statistical significance between two sets of numbers derived from the thirty simulations associated with each of the data sets. The number sets used in the statistical calculations represent the cumulative number of new ideas generated up to a given level of overall responses. The various levels at which number sets were compared were 50, 100, and 250. It is important to note here that the number of simulations used has the potential to affect the significance calculations as the closer the number of samples approaches the total number of responses the more similar the set of responses will be. It was assumed that thirty simulations was sufficiently low to mitigate this possible effect.

Results

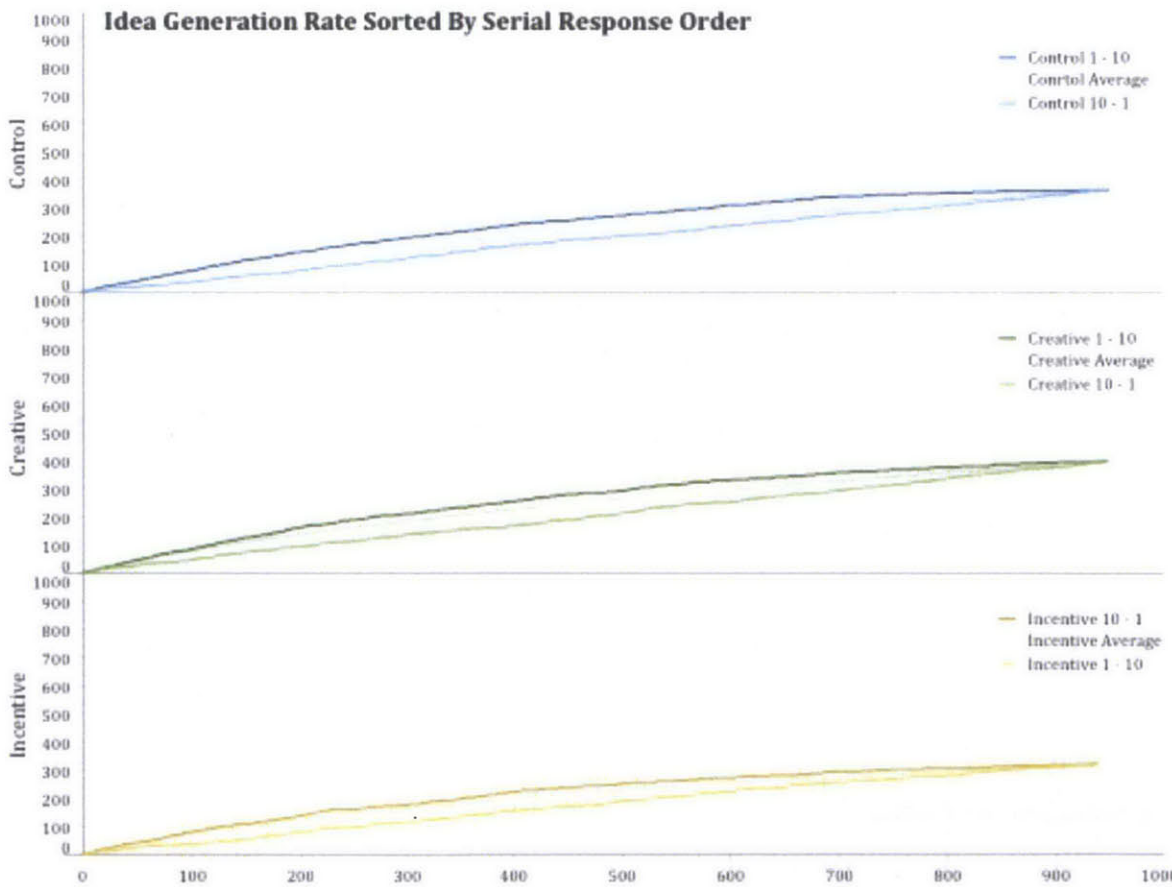
Serial Order Effect

As described in the literature review, the serial order effect expects that fluency will decrease over time and idea quality will increase over time. It is important to note, at this point, that this study differs from previous demonstrations of the serial order effect. Past research that demonstrates this effect using divergent thinking tests that allow respondents to provide as many answers as they can think of in a given amount of time (usually 10 to 15 minutes). The project at hand differs from past studies in that it asked for a discrete number of responses (ten) and allowed workers to complete the task at their own pace with an average HIT completion time was about six minutes and forty seconds. Instead of time stamping the responses, responses were ordered one through ten. The analysis of the serial order effect in this project will use response order as means of measuring the rate of ideas generated over time.

Measuring Quantity

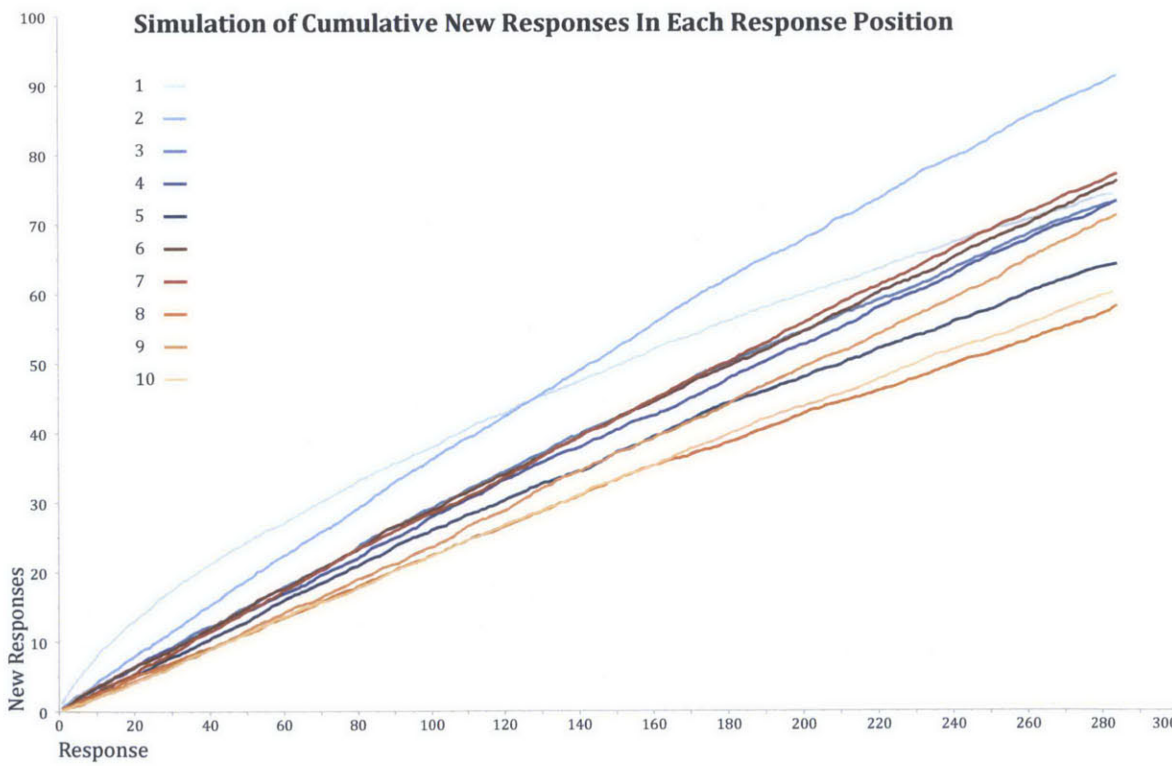
Two different methodologies were applied to analyze the rate of ideas generated in relation to the serial order position of the response. The first technique sorts the responses in ascending and descending serial order in relation to the simulated average rate. Exhibit 7 illustrates this sorting for each of the three question frames. For each case, the charts clearly depicts a more rapid idea generation rate for the descending sorting (10 to 1) when compared to the simulated average. In addition, the ascending sorting (1-10) illustrates a slower, steadier rise when compared to the simulated averages. This indicates that even though the rate of new ideas generated slows over the response period, the novelty of ideas generated increases with time.

Exhibit 7: Idea Generation Rate Sorted by Serial Response Order



The second methodology for measuring the quantity of ideas generated over time involved examining the response generation rate of each of the serial order positions in isolation. To do this, the aggregate pool of responses was sorted to group responses by each of the serial order positions (e.g. all ones, all twos...all tens). Responses were then randomized within each position and the cumulative fluency for each of the response positions was recorded. Thirty simulations were run to calculate an average response rate. This technique demonstrates that when examined in sequential order, the quantity of ideas generated decreases toward the end of the response period. *Exhibit 8* shows these response rates and indicates that more new ideas were generated in the early response positions (1-3) than the later response positions (8-10). This is consistent with previous findings of the serial order effect, demonstrating that the fluency of responses decreases over time.

Exhibit 8: Simulation of Cumulative Fluency for Response Positions



Measuring Quality

In addition to measuring the rate and quantity of ideas generated over time to demonstrate the serial order effect, it is also necessary to measure the quality of ideas generated. For the purposes of this study, the quality of a given response was measured by the relative frequency at which that response occurred within the total pool of responses. This equates idea quality with idea novelty and builds off of past research that bases idea evaluation solely on the measurement of idea novelty (Firestein, 1990; Dennis & Valacich, 1993; Crowne & Ross 1995). Completely unique responses were given a novelty score of 1. Less than completely unique responses were assigned a novelty score based on the following formula:

$$Novelty = 1 - \frac{Total\ Occurances\ of\ Response}{Total\ Responses}$$

For example, *Build House* was the most common response given with a total of 119 occurrences out of 2,840 total responses for a novelty score of

$$Novelty\ of\ Build\ House = 1 - \frac{119}{2{,}840} = 0.958$$

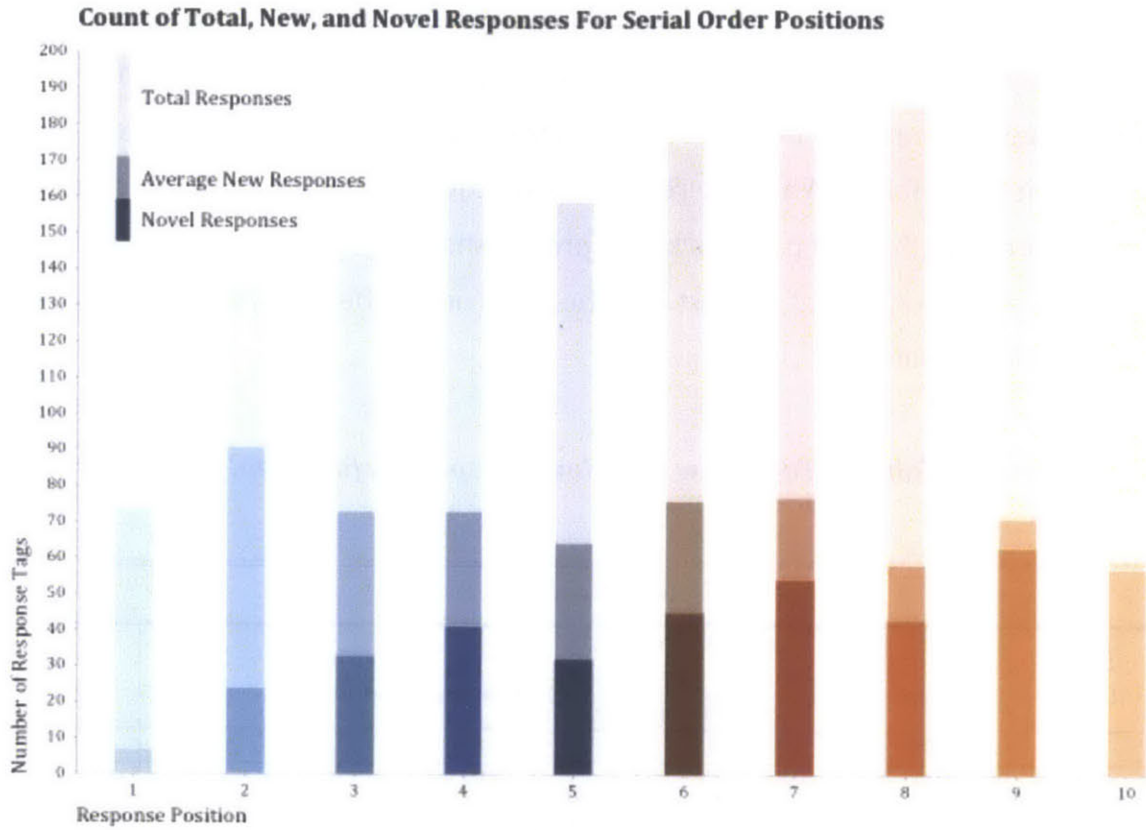
On the other hand, 399 of the 2480 responses were completely unique, occurring just once in the overall set of ideas generated, and received a score of 1.0. This measure of relative frequency stands as the sole metric by which ideas are evaluated in this work.

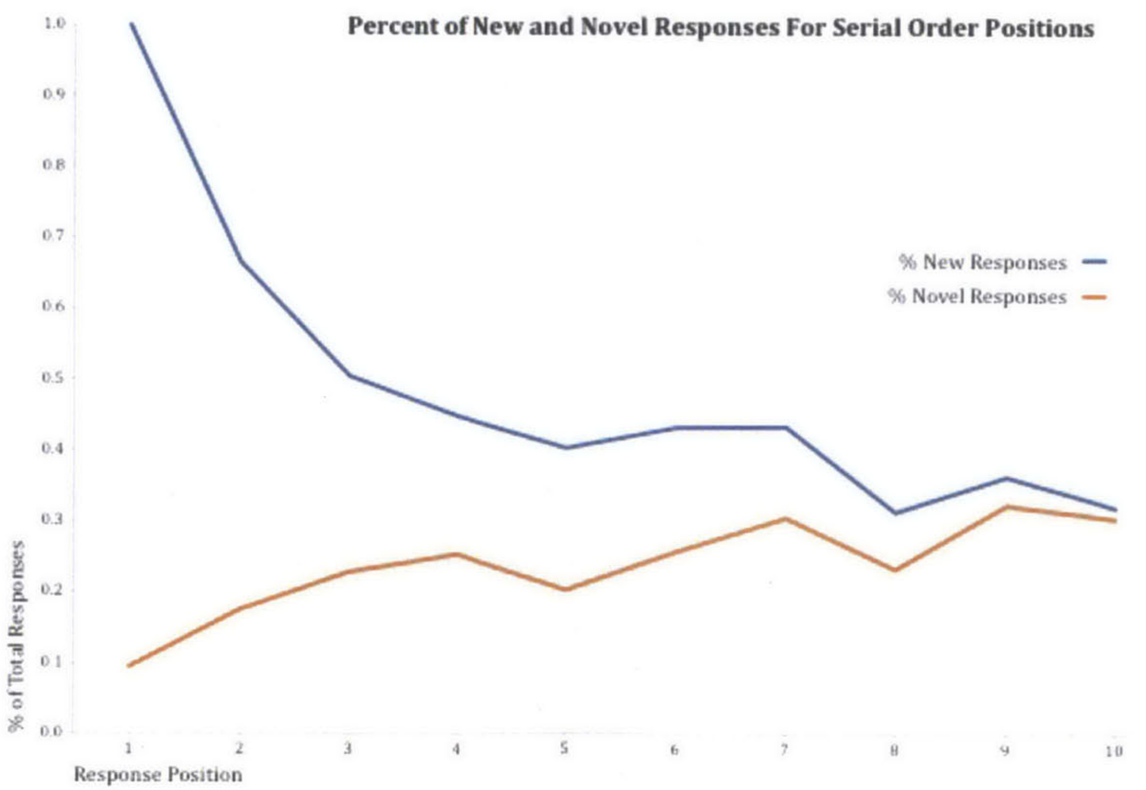
Scoring the novelty of responses in this way allows an average novelty score to be calculated for each of the response positions. Additionally, it identifies each of the completely unique ideas (Table 1). This analysis indicates that responses in the later response positions were generally more novel than responses in the earlier positions. The total number of completely unique ideas in the first half of responses, for example, was 137, as opposed to 262 completely unique responses in the later half. When looked at as a percentage of the number of responses given for each of the response positions, generation of new ideas decreased over time, and the percentage of completely novel ideas, increased over time. These trends are illustrated in Exhibit 9. While the response collection methodologies used in this work differ from those of past work, this project shows several unconventional testing methodologies that suggest that the serial order effect is present in this crowd based population.

Table 1: Total new and Novel Responses for Serial Order Positions

Serial Order Effect							
Serial Order	Number of Responses	# of Response Tags	Average # of New Responses	% New Responses	# of Novel Responses	% Novel Responses	Average Novelty
1	284	74	74	100.0%	7	9.5%	0.979
2	284	137	91	66.4%	24	17.5%	0.985
3	284	145	73	50.3%	33	22.8%	0.990
4	284	163	73	44.8%	41	25.2%	0.992
5	284	159	64	40.3%	32	20.1%	0.992
6	284	176	76	43.2%	45	25.6%	0.993
7	284	178	77	43.3%	54	30.3%	0.997
8	284	186	58	31.2%	43	23.1%	0.993
9	284	196	71	36.2%	63	32.1%	0.995
10	284	188	60	31.9%	57	30.3%	0.994

Exhibit 9: Count of Total new and Novel Responses for Serial Order Positions





Question Frames

Explicit Instructions

In this portion of the analysis, data from the thirty simulations of the question frame with explicit instructions to be creative was compared to data from the control question frame. The hypothesis stated that instructing respondents to think creatively would increase the overall rate of idea generation. The comparison of these two factors is highly significant with a P-value of 0.0087 at the lowest response level and confirms the hypothesis that explicit instructions to think creatively will increase the rate and quantity of ideas generated. Comparisons at higher response levels are more likely to show significance as the number of samples approaches the total number of responses.

Exhibit 10: Cumulative Fluency of Responses to Three Test Question Frames

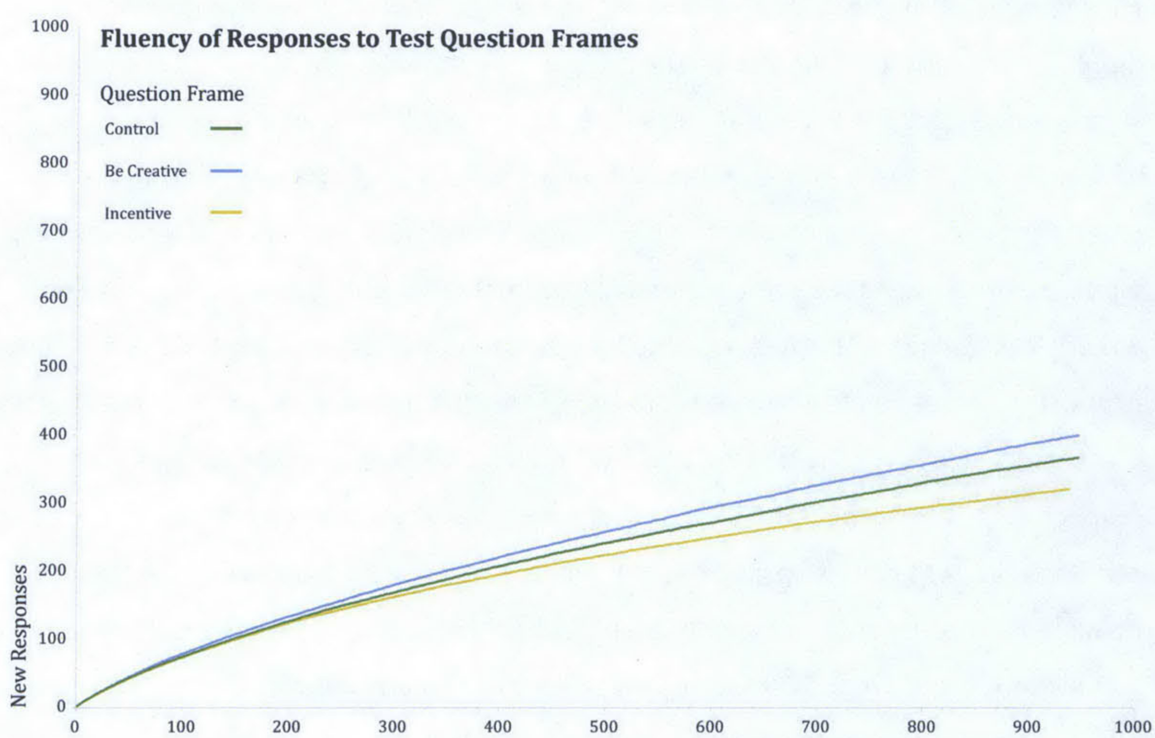


Table 2: Significance Calculation for Hypothesis 2

Hypothesis 2 : Explicit Instructions					
Level	Control Mean	Control SD	Be Creative Mean	Be Creative SD	P
50	41.83	2.51	43.37	1.81	0.0087
100	72.93	4.45	76.77	4.16	0.0011
250	147.60	5.7	155.27	5.97	0.0001

Monetary Incentive

The introduction of an additional monetary reward was hypothesized to increase the response rate. The charts illustrate that, when compared to the control group, respondents generated fewer new ideas when a monetary incentive was provided (above and beyond the reward paid to the worker for completing the task). Statistical analysis shows that comparisons at the lower levels of responses (50, 100) between this group of respondents and the control group is not significant. At the 250-response level, however, the comparison between the datasets is highly significant. There are several factors that may account for this finding, first among which is the fact that respondents were paid a base amount for completing the task, regardless of their level of effort or creativity. The abstract promise of a potentially higher reward may not have been enough to further incentivize a worker who is already being paid. Another possible factor that may effect this analysis is the high proportion of respondents from India (90%) for this question frame, compared with 63.4% and 36.6% for the control and explicit instructions question frames respectively. While the data does not represent a statistically significant relationship between the control group and the incentive group, it does suggest that providing an additional monetary incentive has the potential to reduce creative idea generation. More work is necessary in order to draw any conclusions regarding the effects of monetary incentives on the idea generation rate of crowd-based populations.

Table 3: Significance Calculation for Hypothesis 3

Hypothesis 3 : Monetary Invcentive					
Level	Control Mean	Control SD	Incentive Mean	Incentive SD	P
50	225.00	2.51	41.3	2.6	0.422
100	0.08	4.45	72.33	3.79	0.5761
250	6.31	5.7	142.57	6.44	0.0022

Demographic Sorting

Two levels of analysis were conducted to understand the relationship between two demographic factors (nationality and gender) on the rate of idea generation. The first analysis for each demographic factor uses the aggregated set of responses across all three question frames to test the significance of the relationship. Thirty simulations of this aggregated list were run to generate an average idea generation rate for respondents in each of the demographic groups. The second level of analysis examines the relationship between each of the demographic factors the idea generation rate for each of the three question frames. Thirty summations were run for each of these factors and significance was calculated using a t-test.

Nationality

The first level of analysis (i.e. aggregating all responses across the three question frames), yielded a total 1,740 responses from Indians and 1,010 responses from Americans. Exhibit 11 graphs the fluency rates of these two groups two groups and shows that the group of American respondents was able to generate nearly as many responses as the group of Indian respondents in about 60% as many responses. This further illustrates how the skew in population demographics may be attributable for the less than significant findings in the monetary effects section of the analysis. Sets of numbers from each group of responses were compared against one another, and statistical comparison of these two datasets indicates no significant relationship between the two groups (Table 4).

The second statistical study analyzes the performance of the two groups (Indians and Americans) across the three question frames (Exhibit 12). In this analysis, the set of numbers generated in the thirty simulations of responses from American for the be creative question frame and the incentive question frame were compared to the set of numbers generated from the simulations of American responses to the control group. The analysis indicates a significant relationship between explicit instructions to be creative and an increased idea generation rate for American respondents (Table 5). Furthermore, the analysis shows a significant relationship between an incentive to be creative and a decrease in the rate of ideas generated by the group of respondents from India (Table 6).

Exhibit 11: Fluency of Aggregated Responses By Nationality

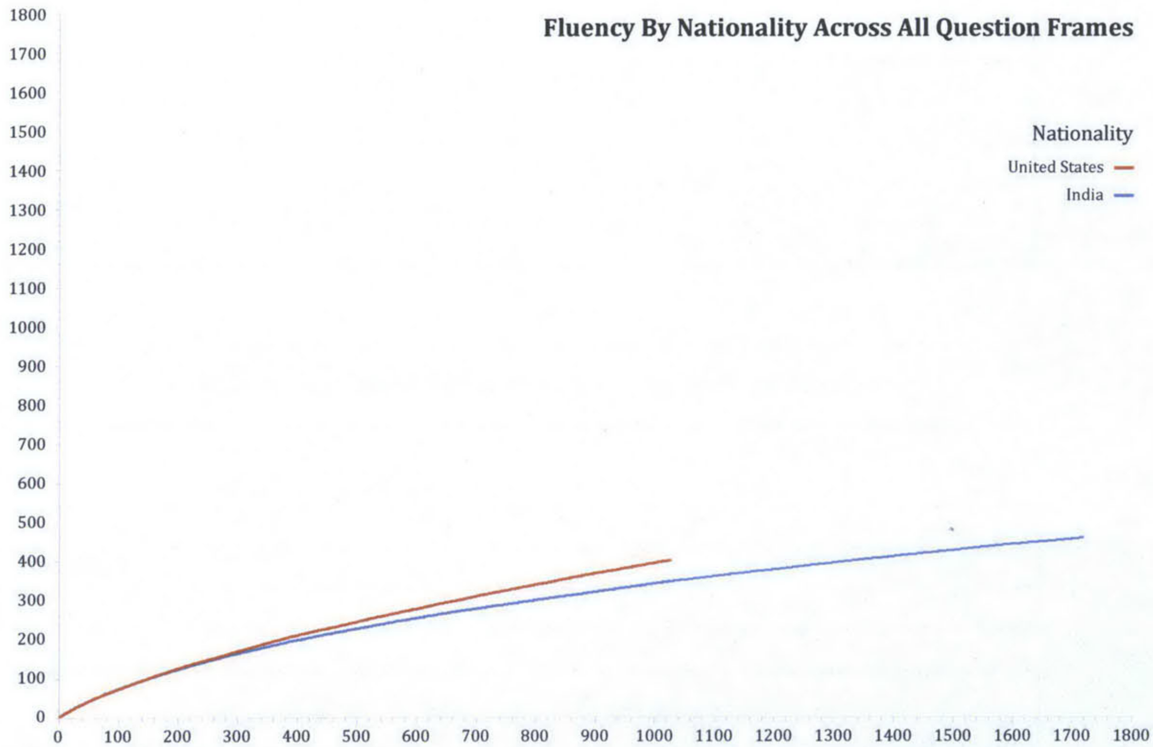


Exhibit 12: Fluency of Responses By Nationality For Three Question Frames

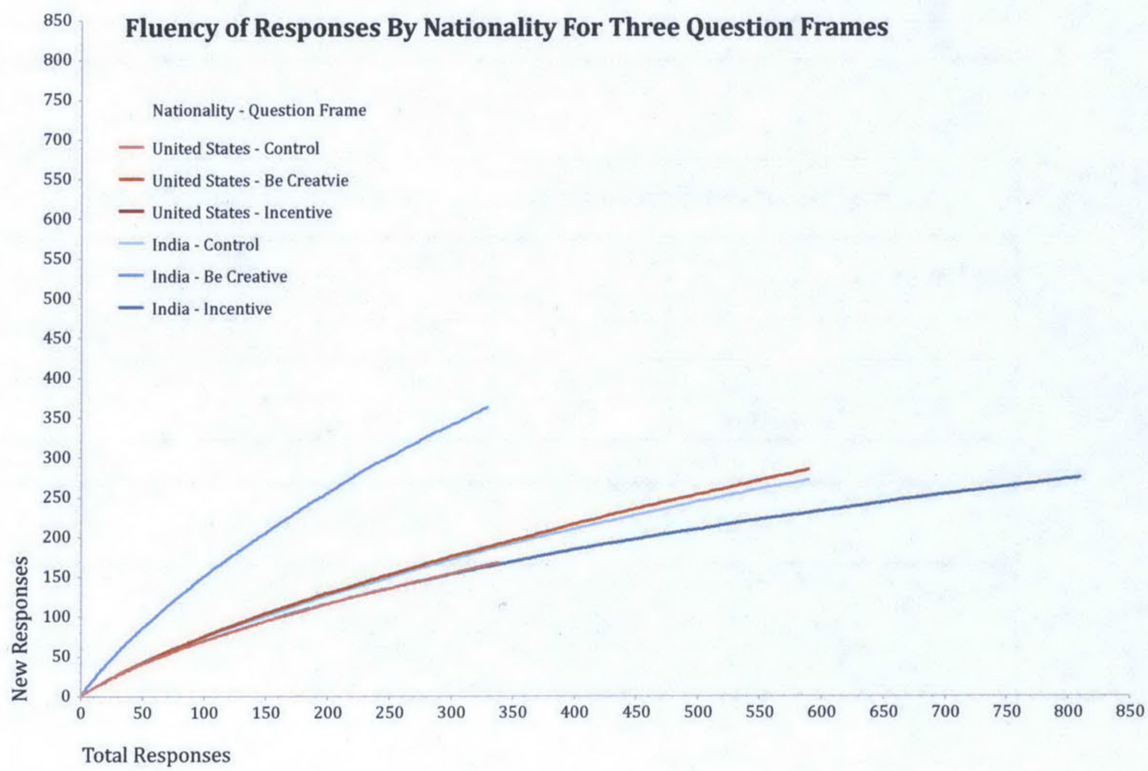


Table 4

Nationality Comparison					
Level	USA Mean	USA SD	India Mean	India SD	P
50	41.10	2.52	41.50	2.43	0.5345
100	71.87	4.07	72.63	3.90	0.4591
200	124.70	4.89	122.27	6.72	0.1141
250	146.67	6.69	143.77	5.66	0.0750

Table 5

Effects of Explicit Instructions on US Respondents					
Level	Control USA Mean	Control USA SD	Creative USA Mean	Creative USA SD	P
25	22.4	1.54	22.7	1.26	0.4137
50	40.1	2.5	41.57	3.08	0.0474
100	69.13	3.79	74.7	4.17	0.0001

Effects of Incentives on US Respondents					
Level	Control USA Mean	Control USA SD	Incentive USA Mean	Incentive USA SD	P
25	22.4	1.54	22.03	1.45	0.3471
50	40.1	2.5	40.6	1.9	0.3867
100	69.13	3.79	-	-	-

Table 6

Effects of Explicit Instructions on Indian Respondents					
Level	Control India Mean	Control India SD	Creative India Mean	Creative India SD	P
25	22.8	1.42	23.03	1.56	0.5481
50	42.03	2.41	42.3	2.6	0.6822
100	73.6	3.9	74.73	4.5	0.3017

Effects of Incentives on Indian Respondents					
Level	Control India Mean	Control India SD	Incentive India Mean	Incentive India SD	P
25	22.8	1.42	22.53	1.22	0.4399
50	42.03	2.41	40.9	2.56	0.0832
100	73.6	3.9	69.17	3.47	0.0001

Gender

Analysis of the aggregated set of responses for females and males across the three question frames yielded a total of 1290 responses for the female population and 1550 responses for the male population. Thirty simulations were run to calculate an average response rate for each group (Exhibit 13). This chart indicates that females had a slighlty higher response rate than males. Statistical analysis shows no significant relationship between these two groups at lower fluency levels (50 and 100). Highly significant results are found when the groups are compared at the 200-response level (Table 7).

The second level of analysis shows high levels of significance for females when comparing the set of responses to the control question versus those for the incentive question frame (Table 8). A similar though less significant relationship is shown for the male population (Table 9). This indicates that the introduction on an incentive into the divergent thinking task decreased the response rate for both males and females. Highly significant results were also found when comparing the response rate of male to the be creative question frame. This indicates that explicit instruction had the effect of increasing the divergent thinking abilities of males in the test population. Fluency charts comparing gender performance are illustrated in Exhibit 14.

Table 7

Females to Males					
Level	Female Mean	Female SD	Male Mean	Male SD	P
50	42.00	2.67	41.77	3.17	0.7587
100	74.53	4.27	74.90	4.36	0.7433
200	152.27	6.47	125.63	4.96	0.0001
250	151.90	6.57	147.50	5.28	0.0059

Exhibit 13: Fluency of Aggregated Responses By Gender

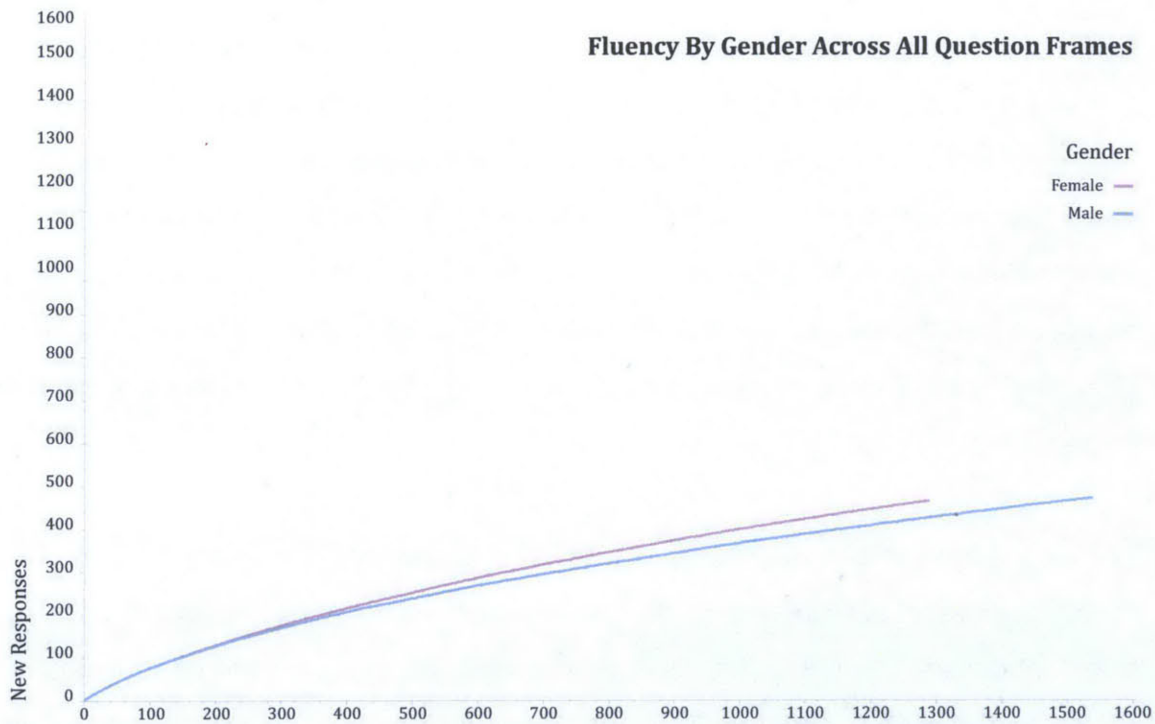


Exhibit 14: Fluency of Responses By Gender For Three Question Frames

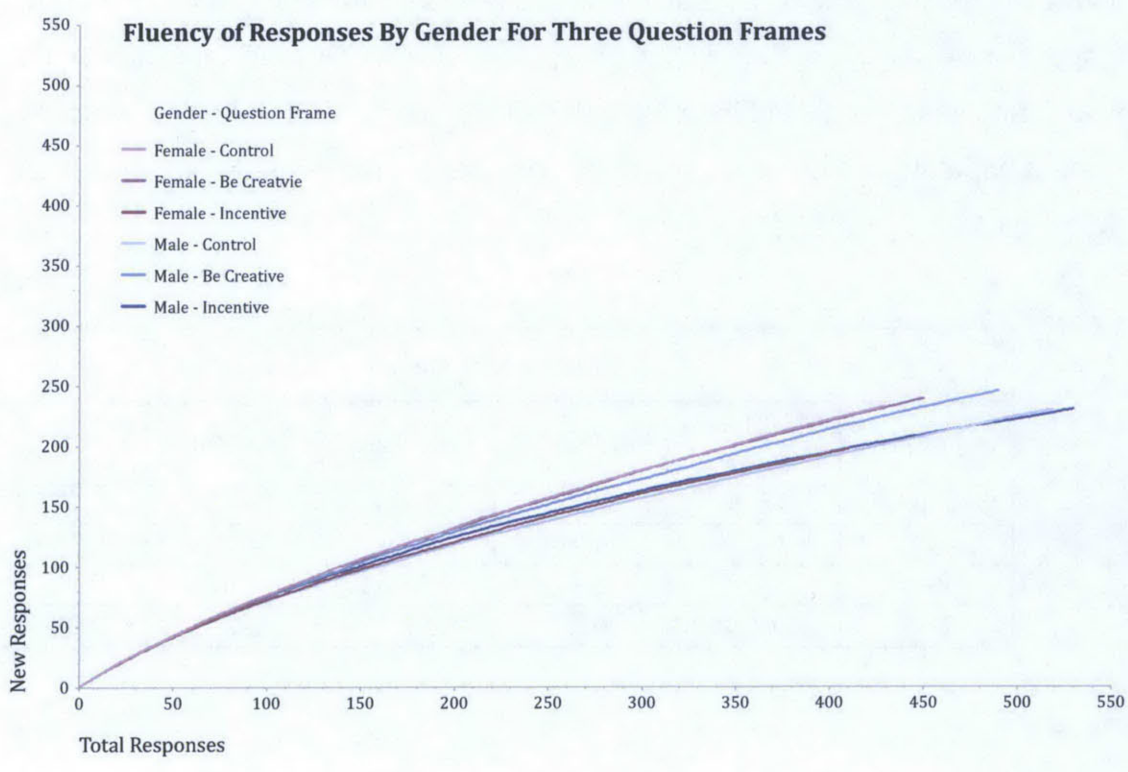


Table 8

Effects of Explicit Instrucions on Female Respondents					
Level	F Control Mean	F Control SD	F Creative Mean	F Creative SD	P
25	23.2	1.3	22.63	1.77	0.1628
50	42.4	2.14	42.27	2.6	0.8291
100	76.17	3.13	75.2	3.22	0.2433
150	105.8	3.96	105.17	3.74	0.5269

Effects of Incentives on Female Respondents					
Level	F Control Mean	F Control SD	F Incentive Mean	F Incentive SD	P
25	23.2	1.3	22.37	1.61	0.0311
50	42.4	2.14	72.07	3.08	0.0001
100	76.17	72.07	3.13	3.08	0.0001
150	105.8	3.96	97.6	4.67	0.0001

Table 9

Effects of Explicit Instrucions on Male Respondents					
Level	M Control Mean	M Control SD	M Creative Mean	M Creative SD	P
25	22.6	1.19	22.8	1.56	0.5794
50	40.67	2.63	41.23	2.54	0.3996
100	71.17	3.54	74.43	3.74	0.001
150	95.57	4.71	102.87	4.75	0.0001

Effects of Incentives on Male Respondents					
Level	M Control Mean	M Control SD	M Incentive	M Incentive	P
25	22.6	1.19	22.2	1.27	0.2135
50	40.67	2.63	41.23	2.54	0.3996
100	71.17	3.54	74.07	2.77	0.0008
150	95.57	4.71	100.03	3.83	0.0002

Limitations and Areas for Future Inquiry

One main limitation of this study stems from the difference in how the rate of responses was measured in this study as opposed to the historical research. Much of the recent work uses a time-stamp of some sort to indicate the exact time a response was created and the duration between responses for a given respondent. Past work also allowed respondents to come up with as many responses as they could think of in a given amount of time. This project differed from past work in two important ways. First, durations between responses were not calculated, rather response position was used as a measure of time. Secondly, respondents were asked to submit a discrete number of responses as opposed to as many as they could think of. This difference meant that all respondents provided the same number of responses. Although the basic AMT interface does not allow for this kind of complexity in the response collection process, future work on divergent thinking tests might be able to design a more accurate way of collecting time-based data.

Certain issues with the AMT interface and response solicitation and collection processes presented limitations for this research. One such imitation stemmed from the fact that multiple HITs were published simultaneously. This allowed some, but not many, workers to provide responses to more than one question frame. While this gave some worker an opportunity to answer the same question more than once, the methodology used for filter respondent data ensured that only one set of responses associated with a Worker ID would be counted for a give question frame. A further complexity of the user and worker interface that may have affected the research stems from the fact that workers can create and use multiple accounts, with different Worker IDs, and have the potential in this way to submit more than one set of responses to a given question frame. This limitation could might not be easily overcome for future research, but still has the potential to skew results.

Although the findings of this work that demonstrate the presence of the serial order effect and the effects of explicit instructions to be creative are consistent with those of much of the previous work aiming to measure and rate quantity and quality, there are several areas where the present work could continue. First and foremost, a more robust evaluative framework could be applied to the scoring of the quality of responses. As opposed to simply counting the number and relative frequency of ideas generated, novelty could be evaluated using the sub-dimensions of originality and paradigm relatedness, as outlined in the work of Dean et al, 2006. This work could be conducted using the existing dataset and might provide more reliable results for the demonstration of the serial order effect. Furthermore, the data generated for this research could be evaluated using the full framework for creative idea evaluation proposed by Dean et al, 2006. Evaluating the responses using more thorough metrics for dimensions of creativity could shed light on the creative potential of the collective intelligence associated with the AMT platform.

Discussion

Measuring the quantity and quality of ideas is a long studied aspect in the field of creativity research. The advent of crowd-based populations provides researchers a new source of potential subjects to test long-held and emerging findings associated with creative idea generation and evaluation. This work confirms one of the oldest empirical findings in the study of creativity by demonstrating the presence of the serial order effect in this crowd-based population. The work also confirms that explicit instructions to think creatively can increase the quantity of responses generated. The work also shows that while there is no significant relationship between monetary incentives the quantity of ideas generated for the overall population, there is a significant that shows a decrease in ideas generated for Indian, female and male respondents when presented with an incentive. The findings of this research suggest that AMT has the potential to be an effective population for testing creative performance for divergent thinking tasks.

The work presented in this paper represents a new and valuable direction for research exploring creativity and creative idea generation. Past research explored how in-person groups, nominal-groups, and individuals generate new ideas. This work applies the constructs developed by past researchers to a new population and explores how massive crowds generate new ideas. This is important because technology is changing the ways in which workers communicate, generate, and synthesize ideas, so creativity is more than just in-person teams collaborating in a workplace setting. Additionally, the data collected from crowd workers could allow researchers the opportunity to explore how group composition might affect the rate and quality of ideas generated. Because the data can be sorted and simulated, various hypothetical group compositions could be tested against one another in order to identify an ideal group composition to maximize performance on a divergent thinking task.

Appendix

Appendix 1

Respondent Education Distribution Across Demographic Measures

	Some High School	High School Graduate	Some College, No Degree	Bachelor Degree	Graduate Degree	
United States	3	6	32	45	15	101
India	13	5	13	109	34	174
Other	0	2	1	3	3	9
18-24	8	5	12	46	4	75
25-34	1	5	22	76	32	136
35-44	3	1	7	22	13	46
45-54	0	0	4	9	2	15
55-64	2	2	1	3	1	9
65+	2	0	0	1	0	3
Total	**16**	**13**	**46**	**157**	**52**	

Appendix 2

Respondent Gender Distribution Across Demographic Measures

	#	% of Total	# F in Range	# M in Range	% of Range F	% of Range M
18-24	75	26.4%	27	48	36.0%	64.0%
25-34	136	47.9%	67	69	49.3%	50.7%
35-44	46	16.2%	22	24	47.8%	52.2%
45-54	15	5.3%	8	7	53.3%	46.7%
55-64	9	3.2%	4	5	44.4%	55.6%
65+	3	1.1%	1	2	33.3%	66.7%
Some High School	16	5.6%	6	10	37.5%	62.5%
High School Graduate	13	4.6%	9	4	69.2%	30.8%
Some College, No Degree	46	16.2%	21	25	45.7%	54.3%
Bachelors Degree	157	55.3%	65	92	41.4%	58.6%
Graduate Degree	52	18.3%	28	24	53.8%	46.2%
United States	101	35.6%	54	47	53.5%	46.5%
India	174	61.3%	71	103	40.8%	59.2%
Other	9	3.2%	4	5	44.4%	55.6%
Total	**284**	**100%**	**129**	**155**	**45.4%**	**54.6%**

Appendix 3
Complete List of Response Tags

Art	Art Avant Garde
	Art Create
	Art Take Picture Of
	Art Mosaic
	Medium of Artistic Criticism
	Art On Brick Sell On eBay
	Art Modeling
	Art Collage
	Art Googly Eyes
	Fire Brick
	Design
	Art On
	Design Amusement Parks
	Art Observe Angles
Play / Games	Game Hopscotch
	Game Jenga
	Game New Sport
	Play
	Play and Arts
	Play Game
	Play Minecraft
	Play Catch
	Play Bomb
	Play Hide And Seek
	Play Kids Smartphone
	Play Imaginary Castle
	Play Fetch
	Play Games
	Play Baseball
	Play Game Capture the Flag
	Play Station
	Play House
	Toy
	Toy Phone
	Toy Remote Control
	Toy Yoyo
	Toy Car
	Doll Brick
	Doll Bed
	Perform Play Three Pigs
	Reenact Scene Homer Alone 2
	Juggle
Carve / Scratch	Brush Teeth
	Toothpaste
	Statue / Sculpture
	Engrave Brick
	Carve Brick
	Carve Design
	Carve for Good Handicraft
	Cave Into Something
	Carve Hole Hide Message
	Carve Hole Hide Key
	Carve Venus Avatar
	Hide Key Under
	Loofah Unfriendly Car
	Scrape Mud Off Shoe Bottom
	Scratch Note On
	Scrubber
	Scratch
	Scrub Body
	Scuff Surface
	Scrape Shoe Bottom
	Scrape Brick
	Pumice Stone
	Wash With
	Clean Penis Mosque
Commodity	Collect
	Fair Swap for Two Half Bricks
	Keep Til You Get Creative
	Wrap Brick as Gift
	Keep
	Give Gift
	Gift for House Construction
	Sell
Fill	Fill Hole
	Fill Broken Window Pane
	Block Hole

	Cover hole
	Block Draft
	Block Hole Fence
	Block Sewage Leak
	Block Dam
	Drain Stopper
Crush / Powder	Crush Crackers
	Crush Egg
	Crusher
	Mortar and Pestle
	Mud
	Sand
	Sand Use With
	Sand Hard
	Powder
	Powder Color
	Powder Color Sand
	Powder Rangoli
	Powder Make Paint
	Powder Adulterate Food
	Powder Make Toy
	Powder Draw Picture
	Powder Extinguish Fire
	Powder Siddah Medicine Stop Bleeding
	Powder Adulterate Kumkum
	Powder Adulterate Chili Powder
	Grind Coffee Beans
	Grind Powder
	Grind Object
	Grind Grain
	Grind and Brush
	Grind Powder Make Thing
	Grind Powder Make Paint
	Plant Soil
Surface	Armrest
	Cutting Board
	Pillow
	Plate
	Pedestal
	Stand
	Stand Art
	Stand Cook
	Stand Hot Poker
	Stand Lamp
	Stand Mobile
	Stand Object
	Stand Object to Break
	Stand Oven On
	Stand Painting
	Stand Plant Pot
	Stand Appliance
	Stand Car After Stealing Tire
	Stand Coffee Pot In Campfire
	Stand Toy
	Stand Transformers
	Stand TV
	Stand Vessel
	Support
	Support Flowers
	Support Furniture
	Support Hen House
	Support Ladder
	Support Object
	Support Picnic Table
	Support Structure
	Support Sump Pump
	Support Table
	Level
	Level Object
	Prop Up Car
Cooking	Cook On
	Cook Egg On
	Cook Meat On
	Cook Vegetables
	Cooking
	Cooking Hearth
	Cooking Hole
	Cook Wrap Foil Flatten Chicken
	Eat
	Eat and Fail

	Eat Lollipops
	Knife
	Outdoor Oven
	Oven / Stove
	Pizza Stone
	Sharpen Tools
	Utensil / Tool
Paint	Color Brick
	Color Skin
	Color With
	Extract Color
	Paint
	Paint Art
	Paint Brick
	Paint Color
	Paint Picture On
	Paint Make Planter
	Paint House Number
	Paint Table Number
	Paint House
	Paint With
	Paint as Wizard of Oz Set
Hold Object	Decorate as Bookend
	Bookend
	Hold Can
	Hold Candle
	Hold Incense
	Hold Object
	Hold Pen / Pencil
	Hold Plant
	Hold Fireworks
	Hold Paper
	Hold Plant In
	Hold Thread
	Vessel
Draw / Write	Brick as Measure
	Draw On
	Draw On Chalk
	Draw On Face
	Draw On Puppet
	Draw On Sharpie
	Draw On Sidewalk
	Draw Right Angles
	Draw With
	Emergency Pencil
	Measure Length
	Ruler
	Mark Wall
	Notepad
	Paint Stencil
	Nameplate
	Write Future Plans
	Write Letters
	Write On
	Write On Road / Pavement
	Write On Wall
	Write With Chalk
	Write With Color Pencil
	Write With Pen
	Write On Chalk Message
	Write Message On
	Write Message On Unique Sign
	Write On Brick
	Write With
Weapon	Annihilate
	Attack
	At Sporting Event
	Bang On Floor Tap Tap
	Beat Hard
	Break Enemy Head
	Beat Criminal
	Bullet Barrier
	Destroy
	Fight a Giant
	Flatten Kitten
	Funeral
	Fly Swatter
	Hammer
	Heavyweight Boxing Gloves
	Hit / Hurt Person

	Hit Brick
	Hunt Small Animals
	Kill Bug
	Kill Moles
	Kill Mouse
	Kill Person
	Kill Snake
	Kill Thing
	Kitten Compactor
	Knock Over Cyclist
	Knock Someone Out
	Pond Nail to Hang Trowel
	Pound Stake
	Pound Chicken for Dinner
	Maim
	Rob Liquor Store
	Projectile
	Threaten
	Ward Off Animals
	Weapon
	Weapon Self Protection
Sit / Step On	Balance Max Bricks On Hand
	Balance Objects On
	Balance On Head
	Booster Seat
	Bench
	Chair
	Jump
	Ladder
	Fashion Shoes
	Romper Room Shoes
	Platform Shoes
	Sandal
	Shoe Lift
	Stool / Seat
	Sit On
	Sit Under
	Stand On Stretch Calves
	Stand On to Balance for Fun
	Step On
	Step On To Reach High
	Step On Look Taller
	Step On Balance
	Step On Exercise
	Step On to Kiss
	Step to Cross Puddles
	Stepping Stone
	Step Stool
	Walk On
	Trip On
	Use As Balance
Warm / Heat	Boil for Warmth
	Burn Brick
	Burn Generate Pollution
	Burn Generate Heat
	Heat Proof Mat
	Heat Shield
	Warm Bed
	Warm Body
	Warm Foot
	Warm Grill
	Freeze / Cooler
	Heating Pad, Fire / Towel
	Retain Heat
Decoration	Centerpiece
	Cover Brick
	Craft
	Decorate Brick
	Decorate Boarder
	Decorate Garden
	Decorate Ornament
	Decoration
	Decoration House Exterior
	Decoration Outdoor
	Decoration Plant Crush
	Decoration Paint
	Decoration Pavement
	Decorate Fish Tank
	Decoration Paint Halloween
	Decorate Color Brick

	Decorate Floor
	Decorate Palace
	Decorate Wall
Weight	Add Weight To Project
	Anchor
	Balance Weight On Head
	Ballast
	Stopper
	Stop Vehicle
	Close Lid
	Close Oven
	Close Trash Can Lid
	Counter Weight
	Door Jamb
	Door Stop
	Door Stop Glaze Fire
	Door Keep Garage Open
	Emergency Brake
	Get Out of Gym Free Card
	Immerse Floating Material
	Lift with Penis
	Flatten Chicken Gill
	Flatten Pancake
	Flatten Object
	Flatten Meat
	Hold Window Open
	Hold Object In Place
	Hold Window Open
	Make Cover
	Make Heavy Bag Drown Person
	Make Resister
	Measure Weight
	Paperweight
	Restrain Dog
	Sandwich Press
	Skull hardness Plate
	Trailer Stopper Rolling Forward
	Tie String Make Scale
	Use As Scale (Size)
	Weigh Down Balloons
	Weigh Down Chicken
	Weigh Down Foot Throw River
	Weigh Down Kite Go To Bathroom
	Weigh Down Tarp
	Weigh Down Vehicle Traction
	Weigh Down Yard Furniture
	Weigh Down Tent Door
	Weigh Down Furniture
	Weigh Down Object
	Weigh Down Dead Body
	Weigh Down Item Scraping
	Weigh Down Sink Object
	Weight
	Weight Lift
	Exercise
	Tone Hand Muscle
	Weight Make Sauerkraut
Throw	Alarm Trigger
	Drop Brick
	Drop Off Ledge
	Drop Test Speed Against Feathers
	Football
	Kick Brick
	Throw As Ball
	Throw at Car
	Throw At Cat
	Throw at Cow
	Throw At Dog
	Throw at Goat
	Throw At Person
	Throw At Wall
	Throw Away
	Throw Brick
	Throw Brick Contest
	Throw In Water
	Throw Though Window
	Throw Through RNC Window
	Throw to Break
	Throw at Object.
	Throw Exercise

	Throw in Fight		Smash Pieces Bat
	Throw Check How Far		Smash Cell Phone
	Throw At Object in Tree		Windshield Tester
	Throw Brick Petrol Fire Burn Building		Break Fire Wood
Break	Break	Build	Block Enemy
	Break Apart See Inside		Build / Construct
	Break Brick		Build / Renovate
	Break Brick Game		Build Apartments
	Break Car Shield		Build Arch
	Break Coconut		Build Asylum
	Break Concrete Fill		Build Auditorium
	Break Gain Entry		Build Barricade
	Break Glass		Build Bathroom
	Break Gravel / Stone		Build Bed
	Break in Half		Build Bird Bath
	Break Head		Build Bird House
	Break Karate		Build Bridge
	Break Nuts		Build Building
	Break Object		Build Cabinet
	Break Open Bottle		Build Cage
	Break Open Lock		Build Canal
	Break Protests Violent Agitations		Build Castle
	Break Up Clumps of Clay		Build Catapult
	Break Peanuts		Build Car Shade
	Break Pecans		Build Ceiling
	Break Walnuts		Build Cemetery
	Break Window		Build Chimney
	Break Windows Get Into Car / House		Build Cheese Table
	Break With Hand		Build Church
	Break Concrete		Build Classroom
	Break Lock		Build Column / Pillar
	Break Teeth		Build Complex
	Break Toy		Build Company
	Break Small Brick Legos		Build Compound Wall
	Emergency Window Opener		Build Dam
	Make More Bricks		Build Desk
	Break With Hammer		Build Dog House
	Car Break		Build Drainage System
	Smaller Bricks		Build Factory
	Smash Pieces Anger Management		Build Fence

	Build Fire Pit		Build Ramp for Skating
	Build Fire Pit Recycle Brick		Build Retaining Wall
	Build Fireplace		Build Restaurant
	Build Fish Tank		Build Room
	Build Floor		Build Roof
	Build Fort		Build Sandbox
	Build Fountain		Build School
	Build Foundation		Build Shade
	Build Furnace		Build Shelf
	Build Functional Object		Build Shelter
	Build Garage		Build Shops
	Build Grill		Build Slab
	Build Hospital		Build Small House Animals Insects
	Build House		Build Speed Bump
	Build Hut		Build Stadium
	Build Hut Cows		Build Stage
	Build House Recycled Bricks		Build Steps
	Build Inverted Pyramid		Build Steps Garden
	Build Kiln		Build Structure
	Build Library		Build Table
	Build Lighthouse		Build Tallest Safest Structure Possible
	Build Mailbox		Build Temple
	Build Marriage Hall		Build Temple Earn Money
	Build Material Cement		Build Theatre
	Build Material		Build Toilet
	Build Material Heavy Duty		Build Tomb
	Build Material Sustainable		Build Tower
	Build Mini House		Build Tub Tree
	Build Mosque		Build Tunnel
	Build Museum		Build Wall
	Build Office		Build Water Tank
	Build Outside Counter		Build Well
	Build Palace		Build Window Box
	Build Pillar / Column		Improve Structure Of
	Build Platform / Stage		Join
	Build Pond		Office Tool
	Build Pool		Paste
	Build Port		Recycle
	Build Rail Station		Repair House
	Build Ramp		Repair Wall

	Stack / Pile
Pave / Mark Spot	Baseball Base
	Boarder
	Cricket Stump
	Even Out Curb to Drive Over
	Floor Tile
	Football Post
	Garden
	Gate
	Headstone Pet
	Herringbone Pattern Garden Path
	Landscape
	Line Garden
	Line Garden Recycled Bricks
	Line Path
	Line Pool
	Line Tree
	Liner
	Mark Boundary
	Mark Crime Scene
	Mark Garden
	Mark Place
	Mark Plant
	Mark Property
	Mark Road
	Mark Spot
	Pave Driveway
	Pave Path
	Pave Path Break Brick
	Pave Patio
	Pave Pedestrian Walk
	Pave Road
	Pave Sidewalk
	Pavement
	Separate Objects
	Spacer
Pet Brick	As Pet Brick
	Feed Lunch
	Name Brick
	Take On Walk

	Teach Play Dead
	Teach Stay Still
Friend Brick	Abuse Brick
	Dance With
	Go Camping
	Go On Road Trip
	Go Sailing
	Hate Brick
	Hug Brick
	Imaginary Friend
	Kiss Brick
	Love Brick
	Show Affection
	Show Off
	Sleep With
	Take Bike Ride
	Watch TV With
	Tickle Make Happy Make Smile
Worship	Meditation
	Take as Self
	Worship
Make Object	Make Balance
	Make Bat
	Make Blocks
	Make Flower Pot
	Make Music
	Make Music Drums
	Make Noise Hit Bricks Together
	Make Pottery
	Make Punching Bag
	Make Set
	Make Swing
Move / Hold	Bury Brick
	Carry Brick
	Cuddle With
	Haul Brick
	Hide Brick
	Hold Brick
	Knock On
	Lay Brick

	Leave Brick Alone
	Pick Up Brick
	Smell Brick
	Stare At
	Taste Brick
Tie / Hang	Hang Brick
	Make Mobile
	Tie Note To Brick
Teach With	Example of Cuboid
	Experiment With
	Learn How Its Made
	Teach With
	Teach Masonry
	Teach Boy to Be man
Displace / Add Water	Toilet Tank Less Water
	Displace Water
	Dissolve Acid Science Fair
	Dissolve Brick
	Dissolve Brick Make Clay
	Dissolve Brick Make Paint
	Pour Water On
	Store Water
	Soak in Water to Carry
Unique	As Cheese
	As Emerald
	As Pooja
	Ask Myself What To Do With It
	Auto Body
	Avoid Birds
	Balance On Pet
	Basking Area for Cold Blooded Animal
	Bats Into Insults
	Brick Holes Measure Spaghetti
	Brick Sandwich
	Calculation
	Compare Quality of Brick
	Compare Weight of Brick
	Compare With Other Stones
	Dig With
	Divide Channel

	Door Hanger
	Factory Raw Material
	Farming
	Fertilizer
	Find End of Rainbow
	Fire Extinguisher
	Fit Pipe
	Fix Underground Leak
	Flyover
	Fun-Size Breeze Block
	Gardening
	Ground Lightening Rod
	Heavy Duty Post Card
	Jam Gas Pedal, Leap From Car, Fake Own Death
	Keep Above Kitchen Article
	Keep Hose Away From Plants
	Knife Blunter
	Land safety
	Magic
	Mail Brick
	Make Brick
	Make Layers
	Make Pike
	Make Place to Sheet
	Manage Brick Business
	Measure Gravity of Moon / Planet
	Measure Spaghetti
	Medicine
	Mouse Pad
	Musical Instrument
	Omam
	Other
	Other Use In House
	Packing Material If You Hate Company
	Practical Joke In Backpack Watch Person Lift
	Plant Flower In Hole
	Poster
	Prevent Leaks Chemical Company
	Refracting Light
	Salt Lick

	Seedling Potter Encourage Root Growth
	Sex Toy for Masochists
	Shield For Trees
	Soak in Water to Carry
	Stave Skull Underpayment
	Take to Philosopher to Ask What It Is Used For
	Telephone Bill Reduction Device
	Test Safety Matches
	Think About What Brick Wants
	Trailer Chuck
	Use It
	Use Wisely
	Waive at Basketball Game
	Wash Brick
	Water Filter

References

Alter, A. L., Oppenheimer, D. M., & Zemla, J. C. (in press). Missing the trees for the forest: A construal level account of the illusion of explanatory depth. *Journal of Personality and Social Psychology*.Chialvo, Dante R. and Mark M. Millonas, 'How Swarms Build Cognitive Maps', *NATO ASI SERIES F COMPUTER AND SYSTEMS SCIENCES*, 144 (1995), 439-39.

Amabile, Teresa M. "The social psychology of creativity: A componential conceptualization." *Journal of personality and social psychology* 45.2 (1983): 357.

Baer, John. "Gender Differences in the Effects of Extrinsic Motivation on Creativity*." *The Journal of Creative Behavior* 32.1

Baer, John, and James C. Kaufman. "Gender differences in creativity." *The Journal of Creative Behavior* 42.2 (2008): 75-105.

Beaty, Roger E., and Paul J. Silvia. "Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks." (2012).

Borst, Wilhelmina Angelina Maria. *Understanding Crowdsourcing: Effects of motivation and rewards on participation and performance in voluntary online activities.* Erasmus University Rotterdam, 2010.

Briggs, Robert O., et al. "Quality as a function of quantity in electronic brainstorming." *System Sciences, 1997, Proceedings of the Thirtieth Hawaii International Conference on.* Vol. 2. IEEE, 1997.

Byrne, Cristina L., Amanda S. Shipman, and Michael D. Mumford. "The effects of forecasting on creative problem-solving: An experimental study." Creativity Research Journal 22.2 (2010): 119-138.

Chen, Chuansheng, et al. "Effects of explicit instruction to "be creative" across domains and cultures." *The Journal of Creative Behavior* 39.2 (2005): 89-110.

Christensen , Paul R., J. P. Guilford, and R.C. Wislon, 'Relations of Creative Responses to Working Time and Instructions', *Journal of Experimental Psychology,* 53 (1957), 82.

Dean, Douglas L. et al., 'Identifying Quality, Novel, and Creative Ideas', *Journal of the Association for Information Systems,* 7 (2006), 646-98.

Dennis, Alan R., and Joseph S. Valacich. "Computer brainstorms: More heads are better than one." *Journal of applied psychology* 78.4 (1993): 531.

Dennis, Alan R., and Joseph S. Valacich. "Group, sub-group, and nominal group idea generation: new rules for a new media?." *Journal of Management* 20.4 (1994): 723-736.

DeRosa, Darleen M., Carter L. Smith, and Donald A. Hantula. "The medium matters: Mining the long-promised merit of group interaction in creative idea generation tasks in a meta-analysis of the electronic group brainstorming literature." *Computers in Human Behavior* 23.3 (2007): 1549-1581.

Diehl, Michael, and Wolfgang Stroebe. "Productivity loss in idea-generating groups: Tracking down the blocking effect." *Journal of personality and social psychology*

61.3 (1991): 392-403.

Eisenberger, Robert, and Michael Selbst. "Does reward increase or decrease creativity?." *Journal of Personality and Social Psychology* 66 (1994): 1116-1116.

Eisenberger, Robert, and Stephen Armeli. "Can salient reward increase creative performance without reducing intrinsic creative interest?." *Journal of personality and social psychology* 72 (1997): 652-663.

Eisenberger, Robert, Stephen Armeli, and Jean Pretz. "Can the promise of reward increase creativity?." *Journal of Personality and Social Psychology* 74 (1998): 714-714.

Eisenberger, Robert, and Linda Rhoades. "Incremental effects of reward on creativity." *Journal of personality and social psychology* 81.4 (2001): 728-741.

Eisenberger, Robert, and Linda Shanock. "Rewards, intrinsic motivation, and creativity: A case study of conceptual and methodological isolation." *Creativity Research Journal* 15.2-3 (2003): 121-130.

Ekstrom, Ruth B. "Cognitive factors: Their identification and replication." *Multivariate Behavioral Research Monographs* (1979).

Eriksson, K., & Simpson, B. (2010). Emotional reactions to losing explain gender differences in entering a risky lottery. *Judgment and Decision Making,* 159–163.

Firestien, Roger L. "Effects of creative problem solving training on communication behaviors in small groups." *Small Group Research* 21.4 (1990): 507-521.

Furnham, Adrian, and Tanya Yazdanpanahi. "Personality differences and group versus individual brainstorming." *Personality and Individual Differences* 19.1 (1995): 73-80.

Guilford, J. P. "Creativity American Psicologist." (1950).

Guilford, J. P. "Guilford test for creativity." *Beverly Hills, California: Sheridan Supply Company. PO Box* 837 (1951).

Harrington, David M. "Effects of explicit instructions to "be creative" on the psychological meaning of divergent thinking test scores1." *Journal of Personality* 43.3 (1975): 434-454.

Hennessey, Beth A., and Teresa M. Amabile. "Reality, intrinsic motivation, and creativity." (1998): 674.

Heylighen, Francis. 'Collective Intelligence and Its Implementation on the Web: Algorithms to Develop a Collective Mental Map', *Computational & Mathematical Organization Theory,* 5 (1999), 253-80.

Heylighen, Francis. "Formulating the Problem of Problem Formulation." *Cybernetics and Systems,* 88 (1988), 949-57.

Janzik, Lars, and Cornelius Herstatt. "Innovation communities: Motivation and incentives for community members to contribute." *Management of Innovation and Technology, 2008. ICMIT 2008. 4th IEEE International Conference on.* IEEE, 2008.

Kachelmeier, Steven J., Bernhard E. Reichert, and Michael G. Williamson. "Measuring and motivating quantity, creativity, or both." *Journal of Accounting Research* 46.2 (2008): 341-373.

Kaisser, Michael and John B. Lowe, 'Creating a Collection of Question Answer Sentence Pairs with Amazon Mechanical Turk', *Proceedings of the Sixth International Language and Evaluation (LERC '08)*.

Kittur, Aniket, Ed H. Chi, and Bongwon Suh, 'Crowdsourcing Using Studies with Mechanical Turk', *Proceedings of the twenty-sixth annual SIGCHI conference on human factors in computing systems,* ACM (2008), 453-56.

Liikkanen, Lassi and et al., 'Time Constraints in Design Idea Generation', *International Conference on Engineering Desing, ICED* (2009).

Linsey, Julie S., et al. "Collaborating to success: an experimental study of group idea generation techniques." *ASME Design Theory and Methodology Conference.* 2005.

Little, Greg and et al., 'Turkit: Tools for Iterative Tasks on Mechanical Turk', *Proceedings of the ACM SIGKDD Workshop on Human Computation, ACM* (2009).

MacCrimmon, Kenneth R. and Christian Wagner, 'Stimulating Ideas through Creative Software', *Management Science,* 40 (1994), 1514-32.

Malone, Thomas and Mark Klein, 'Harnessing Collective Intelligence to Address Global Climate Change', *Innovations: Technology, Governance, and Globalization,* 2 (2007), 15-26.

Malone, Thomas Robert Laubacher, and Chrysanthos Dellarocas, 'Mapping the Genome of Collective Intelligence', (2009).

Mason, W. A., & Watts, D. J. (2009). Financial incentives and the performance of crowds. In Proceedings of the ACM SIGKDD Workshop on Human Computation (pp. 77–85). New York: ACM.

Matud, M. Pilar, Carmen Rodríguez, and Joaquin Grande. "Gender differences in creative thinking." *Personality and individual differences* 43.5 (2007): 1137-1147

McLeod, Poppy Lauretta, and Sharon Alisa Lobel. "The effects of ethnic diversity on idea generation in small groups." *Academy of Management Proceedings.* Vol. 1992. No. 1. Academy of Management, 1992.

Miura, Asako, and Misao Hida. "Synergy between diversity and similarity in group-idea generation." *Small Group Research* 35.5 (2004): 540-564.

Mumford, Michael D. "Something old, something new: Revisiting Guilford's conception of creative problem solving." *Creativity Research Journal* 13.3-4 (2001): 267-276.

Nijstad, Bernard A., Wolfgang Stroebe, and Hein FM Lodewijkx. "Production blocking and idea generation: Does blocking interfere with cognitive processes?." *Journal of Experimental Social Psychology* 39.6 (2003): 531-548

Niu, Weihua, and Dan Liu. "Enhancing creativity: A comparison between effects of an indicative instruction "to be creative" and a more elaborate heuristic instruction on Chinese student creativity." *Psychology of Aesthetics, Creativity, and the Arts* 3.2 (2009): 93.

Offner, Anne K., Thomas J. Kramer, and Joel P. Winter. "The effects of facilitation, recording, and pauses on group brainstorming." *Small Group Research* 27.2 (1996): 283-298.

Osburn, Holly K., and Michael D. Mumford. "Creativity and planning: Training interventions to develop creative problem-solving skills." *Creativity Research Journal* 18.2 (2006): 173-190.

Paulus, Paul. "Groups, Teams, and Creativity: The Creative Potential of Idea-generating Groups." *Applied psychology* 49.2 (2000): 237-262.

Peeters, Jef et al., 'Refined Metrics for Measuring Novelty in Ideation', *Proceedings of the IDMME-Virtual Concept,* 127 (2010).

Ross, Joel et al., 'Who Are the Crowd Workers - Shifting Demographics in Mechanical Turk', *Proceedings of the 28th of the International Conference Extended Abstracts on Human Factors in Computing Systems, ACM* (2010).

Ruch, Floyd L., and William W. Ruch. "Employee aptitude survey." *Los Angeles, CA: Psychological Services* (1980).

Santanen, Eric L., Robert O. Briggs, and G-J. De Vreede. "The cognitive network model of creativity: a new causal model of creativity and a new brainstorming technique." *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on.* IEEE, 2000.

Santanen, Eric L., Robert O. Briggs, and Gert-Jan De Vreede. "Causal relationships in creative problem solving: comparing facilitation interventions for ideation." *Journal of Management Information Systems* 20.4 (2004): 167-198.

Scott, Ginamarie, Lyle E. Leritz, and Michael D. Mumford. "The effectiveness of creativity training: A quantitative review." *Creativity Research Journal* 16.4 (2004): 361-388.

Suri, S., & Watts, D. J. (2011). Cooperation and contagion in Web-based, networked public goods experiments. PLoS One, 6(3), e16836.

Thornburg, Thomas H. "Group size & member diversity influence on creative performance." *The Journal of creative behavior* 25.4 (1991): 324-333.

Toubia, Olivier. "Idea generation, creativity, and incentives." *Marketing Science* 25.5 (2006): 411-425.

Ward, Thomas B., Merryl J. Patterson, and Cynthia M. Sifonis. "The role of specificity and abstraction in creative idea generation." *Creativity Research Journal* 16.1 (2004): 1-9.

Wilson, Robert C. et al., 'Factor Analytic Study of Creative Thinking', *Psychometrika,* 19 (1954), 297-311.

Winter Mason, and Siddharth Suri, 'Conducting Behavioral Research on Amazons Mechincal Turk', *Behavioral Research,* 44 (2012), 1-23.

Wolpert, David H. and Kagan Tumer, 'An Introduction to Collective Intelligence', *arXiv preprint cs/9908014* (1999).

Woolley, Anita Williams, et al. "Evidence for a collective intelligence factor in the performance of human groups." *science* 330.6004 (2010): 686-688.

Vernon, Philip Ewart. "Intelligence and attainment tests." (1960).