Using Metrics of Curation to Evaluate Information-Based Ideation

ANDRUID KERNE and ANDREW M. WEBB, Interface Ecology Lab | Computer Science and Engineering | Texas A&M University
STEVEN M. SMITH, Department of Psychology | Texas A&M University
RHEMA LINDER, NIC LUPFER, YIN QU, JON MOELLER, and
SASHIKANTH DAMARAJU, Interface Ecology Lab | Computer Science and Engineering | Texas A&M University

Evaluating creativity support environments is challenging. Some approaches address people's experiences of creativity. The present method measures creativity, across conditions, in the products that people make.

This research introduces *information-based ideation* (IBI), a paradigm for investigating open-ended tasks and activities in which users develop new ideas. IBI tasks span imagining, planning, and reflecting on a weekend, vacation, outfit, makeover, paper, internship, thesis, design, campaign, crisis response, career, or invention. What products do people create through engagement in IBI? *Curation* of digital media incorporates conceptualization, finding and choosing information objects, annotation, and synthesis. Through engagement in IBI tasks, people create curation products. This article formulates a quantitative methodology for evaluating IBI support tools, building on prior creative cognition research in engineering design to derive a battery of *ideation metrics of curation. Elemental ideation metrics* evaluate creativity within curated found objects. *Holistic ideation metrics* evaluate how a curation puts elements together.

IBI support environments are characterized by their underlying medium of curation. Curation media include lists, such as listicles, and grids, such as the boards of Pinterest.

An in-depth case study investigates *information composition*, an art-based medium representing a curation as a freeform visual semantic connected whole. We raise two creative cognition challenges for IBI. One challenge is overcoming fixation—for instance, when a person gets stuck in a counterproductive mental set. The other challenge is to bridge information visualization's *synthesis gap*, by providing support for connecting findings. To address the challenges, we develop *mixed-initiative information composition* (MI^2C), integrating human curation of information composition with automated agents of information retrieval and visualization.

We hypothesize that MI^2C generates *provocative stimuli* that help users overcome fixation to become more creative on IBI tasks. We hypothesize that MI^2C's integration of curation and visualization bridges the synthesis gap to help users become more creative. To investigate these hypotheses, we apply ideation metrics of curation to interpret results from experiments with 44 and 49 participants.

Categories and Subject Descriptors: H.5.2 [**User Interfaces**]: Evaluation/Methodology; H.5.2 [**User Interfaces**]: Graphical User Interfaces (GUI); Human Centered Computing; [**Human Computer Interaction**]: HCI Design and Evaluation Methods

General Terms: Design, Experimentation, Human Factors, Verification

Additional Key Words and Phrases: Creativity support tools, digital curation, ideation, creative cognition, interactive information, sensemaking, information foraging, information visualization, exploratory search

This work is supported by National Science Foundation grants IIS-074742 and IIS-1247126.
Author's address: Andruid Kerne, Interface Ecology Lab, Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843-3112; emails: {andruid, andrew, rhema, nic, yin, jon, sashikanth}@ecologylab.net; stevesmith@tamu.edu.

DOI: http://dx.doi.org/10.1145/2591677

ACM Reference Format:
Andruid Kerne, Andrew M. Webb, Steven M. Smith, Rhema Linder, Nic Lupfer, Yin Qu, Jon Moeller, and Sashikanth Damaraju. 2014. Using metrics of curation to evaluate information-based ideation. ACM Trans. Comput.-Hum. Interact. 21, 3, Article 14 (May 2014), 48 pages.
DOI: http://dx.doi.org/10.1145/2591677

1. INTRODUCTION

In conjunction with the growth of the big data of digital information, creative innovation has emerged as the single most important factor leading to job creation and economic success [National Academy of Engineering 2005, 2010; Council on Competitiveness 2005]. A crux of innovation is *ideation*—that is, the generation of new ideas. We need interactive environments that help people use information in support of ideation activities. To direct the development of such environments, we need to be able to evaluate them. This is not straightforward. As an alternative to examining people's self-reports of experiences or their social interactions, the present research develops a methodology that evaluates the products that people create when engaged in innovation activities involving information.

Among human work and play involving information, *information-based ideation* (IBI) is defined as activities that address creativity and innovation by generating and developing new ideas. We identify the creative products that result from collecting, organizing, and annotating digital information resources and clippings from them as *curation*s [Rosenbaum 2011]. We examine how various media of curation provide particular affordances for IBI, noting the roles of the medium of curation of elements and of the medium of their assemblage—for instance, how a medium enables putting elements together to form a whole. We use principles from art and cultural theory to demarcate an *information-based ideation paradigm* of curation for creative innovation. We draw from cognitive psychology, engineering design, and information design to formulate a battery of *ideation metrics of curation* to evaluate creativity. We conduct controlled experiments in which we compare IBI support tools by using ideation metrics of curation products to measure participants' creativity across conditions.

The 2005 NSF Workshop on Creativity Support Tools identified the need for mixed qualitative and quantitative evaluation methods [Shneiderman et al. 2005]. The present quantitative methodology is based in *creative cognition* [Finke et al. 1992; Smith et al. 1995; Sternberg 1999]. Creative cognition investigates creative production through empirical studies of individual and interactive effects of cognitive factors, using measures such as ideation metrics. We develop ideation metrics of curation by extending prior applications of creative cognition in the field of engineering design (e.g., Shah et al. [2002]) to HCI.

In an in-depth case study of the IBI paradigm, we use ideation metrics of curation to evaluate the medium of information composition and associated interactive environments. The raison d'être of information composition is to support ideation by promoting creative cognition of relationships among curated clippings and annotations. Composition is an artistic methodology [Cage 1961; Lippard 1972; Kandinsky 1994; Krauss 1998]. *Information composition* affords representing a curated set of clippings as a visual semantic connected whole [Kerne et al. 2008] (Figure 1). Implicit structure uses spatial relationships to present information [Marshall and Shipman 1993]. Information composition extends prior notions of implicit structure to further enable flexible organization and expression, adding emphasis on images and blending (overlap with translucence), along with proximity, size, and color. The implicit structure of composition's holistic medium of assemblage is designed to promote synthesis, in contrast to the lists and grids of individuated visual elements afforded by popular curation media, such as listicles [Quenqua 2013] and Pinterest. Our hybrid art-science approach

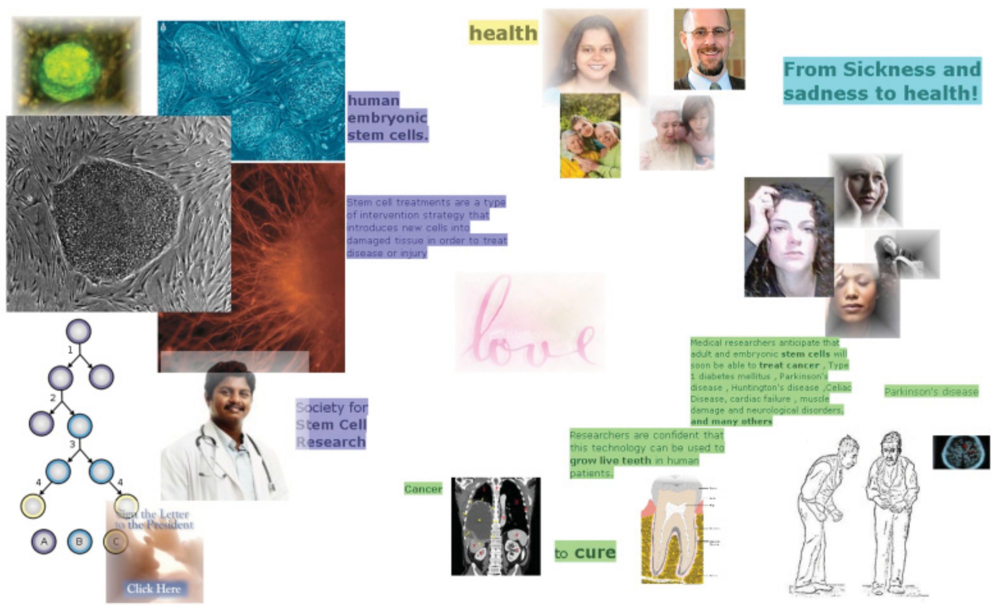

Fig. 1. Information composition curation product for the Stem Cell Research IBI task (Section 3.2), created by a participant in the MI^{2}C Bridges the Synthesis Gap experiment (Section 8), in the Control condition. In contrast to sequential curation media, such as lists or Pinterest, the composition visually synthesizes topics: sickness, sadness, stem cells, treatments, health, and love. We interpret elemental metrics (Section 4.1) using means of this experiment (Section 8.3). $Fluency_{image,text}$ of [19 10] shows that near the mean number of images and fewer than average text clippings were collected. $Flexibility_{document,site,site\_type}$ of $\left[18\ 6\ 3\right]$ represents less diverse than average curation. $Novelty_{image,document,site}$ of $\left[.88\ .84\ .16\right]$ indicates curation of unique documents and images, but from common web sites. Holistic metrics (Section 4.2) are assigned on a 0–3 scale. *Emergence* across raters was 2.5, indicating an original conceptual synthesis, through juxtaposition of stem cells, health, disease, and cure, with the emergent connecting motive of love. *Visual Presentation* is 1 for spatial relationships created using white space. *Relevance* is 3 for an on-topic focus that incorporates implications. *Exposition* of 1.5 (mean among raters) indicates clear statement of only a single theme.

is rooted in triple motives of developing expressive form, generating ambiguous and unexpected stimuli, and providing convivial tools [Illich 2001].

We build this article through three sustained episodes. Episode I: Information-Based Ideation Evaluation Methodology (Sections 2–4) surveys prior work, identifying sensitizing concepts. We specify a range of human activities that constitute IBI. We formulate metrics for evaluating curation products of IBI tasks. *Elemental ideation metrics* evaluate creativity in the digital information objects that people find and curate. *Holistic ideation metrics* evaluate the assemblage of a curation. We derive methods for directly computing elemental metrics. For holistic metrics, which address how elements are put together, we detail procedures for reliable derivation by raters. We prescribe methods for statistical analysis of results.

Episode II: Case Study: Mixed-Initiative Information Composition (Sections 5–8) investigates solutions to two creative cognition problems (Section 5) that IBI support environments face: fixation and the synthesis gap. *Fixation* arises during ideation tasks: a person gets stuck in a counterproductive mental set [Smith and Blankenship 1991]. *Provocative stimuli* help a person overcome fixation [Shah 1998]. Another problem, widely acknowledged in the field of visual analytics, is the need to understand data's meaning in the context of other relevant data [Thomas and Cook 2006]. We call this the *synthesis gap*. To address these creative cognition problems, Section 6 develops principles of *mixed-initiative information composition* (MI^2C), augmenting human authoring of information composition with "automated agents" that iteratively retrieve documents and visualize clippings over time. The two following sections develop studies of MI^2C and apply ideation metrics of curation to show how MI^2C overcomes fixation through provocative stimuli and bridges the synthesis gap.

Episode III: Contributions (Sections 9–11) begins with a discussion of experimental findings, framed by the sensitizing concepts. Next, we develop implications for design of IBI evaluation methodology, and IBI support tools and curation media. We finish by drawing conclusions about the intellectual merit and broader impacts of this research.

EPISODE I: IBI EVALUATION METHODOLOGY

This episode builds an IBI evaluation methodology. We derive sensitizing concepts from prior work across disciplines, connecting art, cognitive psychology, engineering design, and HCI. We sketch a range of IBI tasks and activities. We present procedures for conducting user studies. We formulate metrics for evaluating curation creative products of IBI. We present statistical methods for analysis of resulting data.

2. SENSITIZING CONCEPTS FROM PRIOR WORK

2.1. Creative Cognition

For digital media and interactive systems to support creativity, they should be designed to support the human cognition that produces creative ideas. Mental structures and processes work together to produce creative ideas, discoveries, and innovative products. Although creativity differs across domains, there are commonalities among ways in which creative ideas and discoveries are produced by human cognition [Finke et al. 1992]. Creative cognition studies multiple types of cognitive processes, including ideation, conceptual combination, restructuring, visual synthesis, and visualization [Smith et al. 1995; Sternberg 1999; Kaufman and Sternberg 2010]. The field develops measures that can help HCI researchers understand the effects of media, designs, techniques, and affordances on creativity.

Combinations of dissimilar concepts have been shown to produce emergent properties [Wilkenfeld and Ward 2001]. Consideration of more distant analogies leads to more creative design ideas [Christensen and Schunn 2007]. This supports the hypothesis that the provocative stimuli of unexpected information will increase the likelihood of creative results with novel qualities. Similarly, synthesis of arbitrarily assigned elements was shown to stimulate more creative interpretations than syntheses of intentionally combined units [Finke et al. 1992]. Therefore, automated curation agents, such as those in MI^2C, that select and visualize information that the human user might not intentionally consider could provide an optimal basis for creative visual synthesis.

2.2. Convergent and Divergent Thinking

Intelligence and creativity generally involve convergent and divergent thinking [Guilford 1956; Buxton 2007]. Convergent thinking reduces a space of possibilities to a single answer. Divergent thinking assembles many correct and relevant answers. Divergent thinking is essential to creativity [Guilford 1956; Runco and Albert 1985], problem discovery [Runco and Okuda 1988], and project-based learning [Dym et al. 2005].

2.3. Engineering Design Ideation Metrics

This article extends prior work in engineering design to formulate ideation metrics for evaluating curations that people create through engagement in divergent thinking tasks in which information plays a support role. Human performance on divergent thinking tasks is more difficult to assess than that on convergent thinking [Finke et al. 1992]. Building on the factors of creativity of Guilford [1950], engineering design researchers have measured ideation in solutions to design problems in terms of their fluency (quantity), flexibility (variety), novelty, and quality [Shah et al. 2002, 2003; Song and Agogino 2004; Dym et al. 2005; Linsey et al. 2005]:

—*Fluency* is the number of ideas. According to Darwinian theories of ideation, the more ideas a person considers, through survival of the fittest, the more likely it is that one idea will survive and grow to achieve creativity [Guilford 1950; Simonton 1999].
—*Flexibility/Variety* addresses the investigation of alternative interpretations. It is measured as the number of categories of ideas. Variety provides opportunities for more remote associations, and more remote analogies, both of which are likely to lead to creative products [Mednick 1962].
—*Novelty* is the rareness of an idea. It can be measured as the statistical infrequency of ideas, which requires an appropriate norm for the space of possible ideas [Shah et al. 2002]. Although potentially difficult to assess globally, novelty is straightforward to measure in the context of any particular laboratory experiment. Build a master list and inverted index of all ideas generated by all participants. Then, count the number of participants that presented each idea. The lower the count, the higher the novelty. Thus, the Novelty metric can be seen as analogous to information retrieval's inverse document frequency (IDF) measure [Salton and Buckley 1988].
—*Quality* addresses more contextual features of creative products. Quality is determined by the consensus of expert raters [Shah et al. 2002]. It requires definition of a set of criteria that the raters can independently and consistently apply.

2.4. Evaluation Methods: Creativity Support

Prior creativity support tools have been evaluated in several ways. To convey a sense of this diversity, we summarize representative examples. Kim and Maher [2005] used protocol analysis to compare graphical and tangible user interfaces for collaborative design. They found the tangible user interface improved engagement in spatial cognition. Perer and Shneiderman [2008] used case studies to validate a tool for social network visualization through application to significant real-world problems. TeamStorm, which developed zoomable private and public workspaces for collaborative design, was evaluated by observing and interviewing teams of users [Hailpern et al. 2007]. Carroll et al. [2009] developed the Creativity Support Index, a survey instrument for factor-based self-assessment of creativity in user experiences. Tripathi and Burleson [2012] built a recognizer for reported team creativity based on movement and face-to-face interaction.

Klemmer's group takes an approach more like our own, using features of Web pages to build a model that serves as a basis for traversing associational links to identify related information and present it to users engaged in creative tasks [Kumar et al. 2011; Ritchie et al. 2011]. They report that users of their d.tour software produced more domain-diverse examples than those who used search engines [Ritchie et al. 2011]. This finding is constituted along the lines of the *Flexibility/Variety* ideation metric (see Section 2.3). However, they did not explain how to measure Variety or situate it amidst a battery of ideation metrics of curation.

Dow et al. [2010] presented a novel measure of creativity support: posting graphics created by study participants as web ads and measuring click-through. They measured Variety by having crowdsourced workers engage in pairwise rating of combinations of ads created by other study participants. Like ourselves, they observed that the judgements of raters were not necessarily "calibrated"—that is, that they could lack consistency. For elemental ideation metrics, the present research develops computational methods, eliminating dependence on raters. For holistic metrics, we clearly articulate rating criteria and procedures, resulting in a high level of inter-rater consistency.

2.5. Representing Digital Found Objects in Media of Curation

The present research develops methods for assessing curation tools when the user's work involves ideation. *Digital curation* means the process of forming a conceptual focus, choosing information resources, and developing ideas by combining chosen information [Rosenbaum 2011]. Digital curation involves searching, browsing, reading, reflecting, collecting, organizing, annotating, and expressing. We build a framework for analysis of digital media of curation, building on the technique of collecting and presenting found objects that emerged in conceptual art during the 20th century.

2.5.1. Digital Found Objects. The first concrete step of curation is to *choose* information resources to collect. Dada artist Marcel Duchamp's inception of conceptual art articulated the technique of collecting an object as a creative act. It began in 1917 with *Fountain*, a urinal that he labeled as art and submitted, under the pseudonym R.MUTT, to the Society of Independent Artists exhibition [Lippard 1972]. For Duchamp, *Fountain* was an instance of a *readymade*—that is, a found object that he could collect and recontextualize by creatively presenting it in a new context. Duchamp said, "Whether Mr. Mutt with his own hands made the fountain or not has no importance. He CHOSE it." As part of entering *Fountain* in the exhibition, Duchamp authored metadata for it. He created a new reading of *Fountain*—that is, a new meaning—and a new sense of the process of creating work for presentation. The act of collecting and presenting by choosing, labeling, and recontextualizing was thus formulated as an art-making technique.

Duchamp's methodology is alive and well on today's Internet. People find information resources on the Web. They collect these digital found objects—that is, digital readymades. The meanings and impacts of digital readymades are transformed through contexts of curation. Extending Duchamp's prescient insight on the creativity of choosing readymades, digital curation sites are transforming the media landscape. Pinterest provides a social curation authoring and presentation environment, in which people collect and share digital found objects on personal topics such as "Have Love Will Travel" [Maccariello 2013]. Blogs like `huffingtonpost.com` and `tumblr.com` transform the production of news through blogging and re-blogging. Microblogging in Twitter and Facebook function similarly. The popularity of these redigested forms of media results in reduction of traditional in-the-field news gathering [Utley 1997; Babb 2012].

One way that people curate and repurpose Web content is through "listicles," articles comprising a title, short description, and an annotated list of digital found objects. On `buzzfeed.com`, the inauguration of the president of the United States was rendered as listicles: "24 Delightful Inauguration Firsts" [Johnson 2013], "23 Reasons Sasha and Malia Stole the Inauguration" [Yapalater and Notopoulos 2013], and "The 22 Most Fabulous Beyonce Moments from the Inauguration" [Yapalater 2013]. Each of these visually coherent single topic listicles contains clipping found objects curated from a variety of Web site sources (see Section 4.1.2.). Duchamp's methods of choosing what to collect, and of reframing, become reified in systems, sites, and practices of digital found object curation. In social media, people recontextualize digital readymades!

2.5.2. Media of Curation. How a medium of curation enables representation is characterized on two levels: the representation of individual elements and that of the assemblage of the whole. The medium of elements defines how individual underlying digital objects can appear and be interacted with. The medium of assemblage defines the format and method through which elements can be put together and combined.

The medium of a curated element could simply be a found object's metadata (e.g., title, authors, URL), or, as in Pinterest, listicles, and information composition, metadata enhancing material clipped from a source document, to more precisely represent what in that source document matters to the curator, in the context of an IBI activity. Curation thus involves *clipping*, defined as the finding and selection of images and passages from an information source. We investigate linked clippings as an elemental medium of curation that affords illustrating an idea through digital found objects.

To further develop and articulate ideas, curation involves *annotation*: explaining the role of collected resources and their connections, and provoking reflection. Social metadata includes annotations by the curator, friends, and beyond, across the cloud.

Curation further involves the act of assembling a whole creative product: *conceptualization, integration, juxtaposition, combination, exposition, design, and synthesis*. We analyze popular examples of media of curation. Both listicles and Pinterest utilize visual semantic clippings as the medium of the elements of curation. They differ with regard to the medium of assemblage. The linear lists of listicles are easily read on mobile devices and thus have grown popular for sharing in social media like Facebook. Pinterest represents clipping collections in grid layouts called *boards*, enforcing homogeneous width and variable height. Pinterest has pioneered social aspects of curation, with mechanisms such as "repining," which highlights when users collect from each other. The medium of the Pin element affords easy sharing of a single found object. The flow layout of the board assemblage affords reformatting for presentation on devices of varying size and resolution. Typical Pinterest users engage in everyday forms of IBI [Linder et al. 2014].

Pinterest boards, like information composition, constitute a 2D medium for curating bookmarks as visual semantic clippings. However, information composition provides a flexibility of representation that Pinterest lacks. For elements of curation, composition directly supports text clippings; Pinterest supports only image clippings. For the medium of assemblage, Pinterest's grid is fixed, presenting elements only in the order collected by users. Information composition affords freeform design. Composition enables open spatial organization, blending elements, text clippings, and first-class annotations. NB, the novelty ideation metrics developed herein (see Section 4.1.3.) seem opposite to the popularity of a Pin. Relationships between popularity and novelty in social media are a topic for future research.

A study of personal information management asked, "Once found, what then?" [Jones et al. 2002]. The popularity of Pinterest, the site that grew fastest to 10 million monthly users [Chocano 2012], shows how important this question becomes. Representational medium and social engagement impact the curation of digital found objects.

2.6. Paradigms: How People Work with Interactive Information

Paradigms for investigation of how people work with interactive information include sensemaking, foraging, and exploratory search. Belkin et al. [1982] observed that information needs of users engaged in research evolve. Russell et al. [1993] investigated sensemaking, which involves integrating information from multiple sources. Sensemaking is said to play a role in information processing tasks rather than in creating new knowledge, per se. Representations were found to play a key role. People were said to shift representations to reduce the costs of task operations, but evidence for this was not provided. Information foraging theory similarly posited that people meet information needs by optimizing energy gained versus time spent [Pirolli and Card 1999]. Exploratory search addresses search experiences in which information needs are iteratively refined [White et al. 2006]. It is characterized by combining querying and browsing strategies to foster learning and investigation. IBI includes evolving goals, sensemaking, foraging, and exploratory search, yet shifts focus to how people generate and develop new ideas as they work and play with information.

3. IBI TASKS AND EXPERIMENTAL PROCEDURE

We define the scope of IBI tasks. We use this general picture of IBI to develop example laboratory study tasks and instructions. We present a procedure for how to organize IBI support tools experiments, arranging study sessions to mitigate participants' unfamiliarity with new technologies. The underlying goal of the procedure is to align conditions in order to produce consistent, unbiased comparisons of support tools.

3.1. IBI Tasks and Activities

IBI spans personal, academic, career, and organizational human activities, which encompass imagining, conceptualizing, planning, developing, reflecting, and reassuring. Information plays an essential role. IBI tasks go beyond merely understanding prior facts. Thus, although IBI is engaged across life and work, many tasks involving information, from fact finding to basic assessment, are not IBI.

IBI activities connect analysis and synthesis. Goals include to think, get a sense, have ideas, develop interpretations, achieve insight, gain a new vantage, build perspective, and contribute innovation. Personal IBI activities may involve a meal, outing, vacation, makeover, home furnishing, creative medium, relationship, child, transition, or life stage. Academic IBI task areas include developing a paper, internship, education, degree, thesis, or career. Professional IBI task areas include campaigns, exploration, design, art, and invention. Fields in which IBI is important include science, engineering, humanities, arts, architecture, entertainment, business, exploration, and crisis response.

IBI broadly includes individuals' personal experiences of creativity, scaling up to tasks and activities of wider significance. The only requirement for engagement in IBI is that new meanings and ideas are created for the individual, although IBI support tools also address larger scales of classroom, community, organization, and global significance. As such, IBI encompasses what are known as "mini-c" everyday, "little-c" personal, and "big-C" eminent forms of creativity [Beghetto and Kaufman 2007].

3.2. Example IBI Tasks for Laboratory Studies

IBI tasks for laboratory studies are defined so that participants can author meaningful curations in a short amount of time. For within-subjects experiments, we designed a pair of straightforward IBI curation tasks, suitable for undergraduate students, and of similar complexity. The intended time for each task was around half an hour.

3.2.1. The Stem Cell Research IBI Task. This task was employed in both of the following experiments (Sections 7 and 8). Figure 1 shows one participant's curation product.

> Consider the possibilities of Stem Cell Research. Think of new ideas for how it can be used.

3.2.2. The Liberty and Security IBI Task. This task was employed in the following case study only in the within-subjects Synthesis Gap study (Section 8):

> Compare and contrast ideas of liberty and security. Which is more important to you? How do policies and actions of government impact these ideas?

To encourage creativity, each task definition instructed participants to:

—Express your ideas in an information composition.
—Articulate motivations and implications for yourself and society.
—Develop and connect ideas, opinions, and explanations.
—Find and collect a variety of relevant information.
—Use your own ideas to create a coherent composition that develops original themes, strategies, and directions.

—Use the curation to visualize relationships among information elements.
—Connect elements visually and conceptually.
—Develop a clear sense of what you are presenting and why.
—At the same time, let unexpected ideas emerge.
—Create concise annotations to articulate themes and explain what makes the collected elements important.
—Use your time well. Remember to spend some of your time collecting relevant information (seeding ideas), some organizing and designing (to show), and some annotating (to explain).

3.3. Experimental Procedure: Training and Data Collection Sessions

We develop a procedure for running IBI experiments. We include teaching participants how to use new tools. In the studies that follow, participants had no prior experience with the curation medium of information composition or with the particular IBI support tools. To align experimental conditions, experiments were conducted across two sessions: a Training Session and a Data Collection Session. Each lasted up to 1 hour.

The Training Session begins with an instructional video, explaining information composition as a curation medium, and how to use the IBI support tools corresponding to the experimental conditions. Inasmuch as the exploration and curation interface independent variable is constituted differently in different experiments, instructional videos must be customized for each interface. These videos also explain to participants the evaluation criteria that will be used to evaluate their creative products. The criteria are reiterated in condensed form as part of the instructions when ideation tasks are presented and conducted. The video ensures that IBI support tools and evaluation criteria are consistently presented to participants. After watching the video, participants perform practice IBI tasks. Data is not collected or analyzed.

The Data Collection Session is typically conducted around 2 days later. This break between sessions is designed to alleviate mental fatigue from performing IBI tasks and to reduce fixation carried over from the practice task. The break gives participants the opportunity to engage in incubation (Section 5.1) regarding their understanding of the new technology's capabilities and potential.

4. METRICS OF CURATION FOR MEASURING IBI

We build a creative cognition methodology for evaluating creativity support environments people use in performance of IBI tasks. Particular IBI environments and their concomitant media of curation provide experimental conditions-that is, independent variables. We examine the creative products, the curations that people make in performance of the IBI tasks. We develop a set of distinct ideation metrics, forming a feature space of dependent variables with which to evaluate the creativity of curation products. Our ideation metrics descend from the seminal factors of creativity of Guilford [1950], which include fluency, flexibility and novelty in generating ideas, and the synthesis and reorganization of information. Figures 1, 2, and 5 show example curation products of IBI, with associated metrics. To assess the efficacy of environments, we compare results across conditions using statistical tests of significance.

As we analyzed media of curation at two levels, the medium of the elements and that of their assemblage, so we develop two levels of ideation metrics of curation: *elemental ideation metrics*, for assessing creativity within the digital objects collected to fulfill an ideation task, and *holistic ideation metrics*, for assessing creativity manifested through how those elements are put together. We formulate computational methods for computing elemental ideation metrics. Holistic metrics are assessed by independent raters. We articulate methods to consistently operationalize holistic metrics, including procedures for raters to conduct and criteria for each metric.

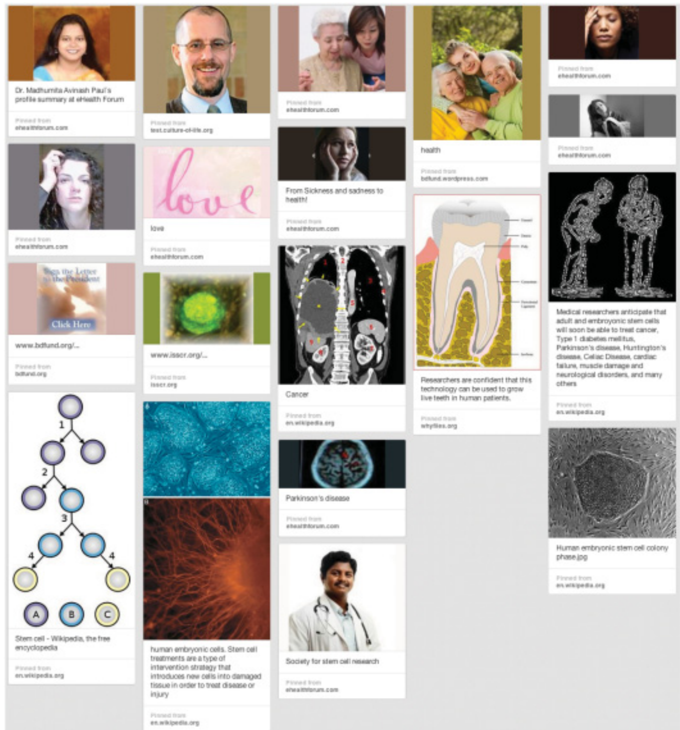

Fig. 2. The same Stem Cell IBI task curation product as shown in Figure 1, re-represented in the medium of a Pinterest board, as an exercise to show how the ideation metrics work with a different curation medium. Since the Pinterest medium does not afford text clippings, text from Figure 1 has been added to image clippings as descriptions to Pins. The elemental metrics for images are the same as for the information composition rendering. Thus, $Fluency_{image,text}$ is [19 0]. Variety and novelty statistics change, as a result of the elimination of text clippings and their source documents. $Flexibility_{document,site,site\_type}$ is [17 6 3]. $Novelty_{image,document,site}$ is [.88 .68 .21]. The change of medium impacts some holistic measures. Emergence is reduced to 1 based on spatial grouping of people and biology images. Visual Presentation is 1, granted for the grid. Relevance is unchanged at 3. Exposition is harder to read in Pinterest, but all of the same text is present, so the same mean 1.5 score again results.

4.1. Elemental Ideation Metrics of Curation

We develop a method for measuring creativity in the digital objects that people collect. We derive elemental ideation metrics of curation from those previously developed in engineering design by Shah et al. [2002, 2003], who treated responses to a design ideation task only as atomized alternatives, not as an assembled whole. Like Shah et al., we compute distinct elemental ideation metrics for Fluency, Flexibility, and Novelty.

We base our application of the prior methods on the observation that each element in a curation is a digital found object, $o_{m,i}$, which consists of a clipping in the medium, m. In the present research,[1] we define the set of media of curated clippings, M, as

$$M = \{image, text\}; m \in M. \tag{1}$$

[1]Future research in metrics of curation can address additional media, such as audio and video clippings.

Since participants curate from Web pages on the Internet, we can express clipping found objects in each of these media m as

$$o_{image,i} = [image\_url\ source\_url] \tag{2}$$

$$o_{text,i} = [text\_string\ source\_url]. \tag{3}$$

Then, we can express each curation, c, as consisting of a set of clipping found objects:

$$c = \{o_{m,0}, o_{m,1}, \ldots\}; \forall m \in M. \tag{4}$$

Any particular study develops a corpus, C, of n curations created by its participants:

$$C = \{c_0, c_1, \ldots c_{n-1}\}. \tag{5}$$

4.1.1. Fluency. Fluency is the quantity, or total number, of ideas generated. It is computed based on the number of digital found objects in each curation. To measure Fluency of a curation c, count how many found clippings the participant has collected. In the present research, we separately compute Fluency for each medium of found object clipping, resulting in a two-element $Fluency_M$ vector, $Fluency_{image,text}$. Thus, for each curation c in C, we compute:

$$Fluency_m(c) = ||o_{m,i} \in c||; m \in M. \tag{6}$$

Each element in $Fluency_{image,text}$ is an integer on $[0, FluencyMax(C)]$. $FluencyMax(C)$ is the largest number of clipping readymades for each medium collected among all curations created by all participants in a study.

Note that a high level of Fluency, while valuable, also presents challenges. Generating more ideas and finding more objects give you more chances to come up with great ones. When you get to presentation and synthesis of the whole, high Fluency also means higher cognitive load: more to look at, think about (working memory), and arrange. This issue is not addressed by Shah et al. because they do not engage participants in working with the entire creative product for an ideation task. In consideration, we will investigate potential negative correlations (Section 9.6) between increased Fluency, and holistic metrics such as Relevance and Visual Presentation (Section 4.2.2).

4.1.2. Flexibility/Variety. Flexibility/Variety involves consideration of alternative interpretations, which means ways of thinking and viewpoints. Flexibility in thinking describes the cognitive process of trying out a Variety of different ways of looking at a problem. Flexibility measures the span of the solution space explored during ideation. There can be many ways to constitute distinct aspects of Flexibility (see Section 10.1.4). We presently construe Flexibility as the diversity of the information sources from which digital found objects are curated.

Flexibility is measured at three levels of information source granularity, g,

$$G_{Flexibility} = \{document, site, site\_type\}; g \in G_{Flexibility} \tag{7}$$

producing the three-element $Flexibility_G$ vector, $Flexibility_{document,site,site\_type}$.

The $Flexibility_{document}$ metric measures the diversity of source Web pages from which clippings are curated. To compute, form the set of all unique source document locations, $SourceUrls(c)$, for the clipping found objects in a curation c:

$$SourceUrls(c) = \bigcup_{o \in c; m \in M} o_m[source\_url] \tag{8}$$

$$Flexibility_{source}(c) = ||SourceUrls(c)||. \tag{9}$$

$Flexibility_{site}$ measures the diversity of Web sites, by domain, from which clipping found objects are curated. To compute this metric, define a $WebSites$ operator, which operates on a *url*. *WebSites*(*url*) computes the top-level Web domain, not just the host name. For example, both `money.cnn.com` and `weather.cnn.com` reduce to the single domain `cnn.com`. To compute, form the set of all unique source Web sites, *SourceSites*(*c*), from the source URLs of the clipping found objects in curation *c*:

$$SourceSites(c) = \bigcup_{url \in SourceUrls(c)} WebSites(url) \tag{10}$$

$$Flexibility_{site}(c) = ||SourceSites(c)||. \tag{11}$$

$Flexibility_{site\_type}$ measures the diversity of kinds of Web sites (i.e., the number of unique categories of Web site) among the found objects curated in *c*. This is accomplished by grouping together types of sites such as news/media, blogs, and portals. The OpenDNS community categorizes domains by crowdsourcing the labeling of Web sites, producing a comprehensive dataset [OpenDNS 2013]. We use this dataset to define a *SiteTypes* operator, which operates on a Web site, *s*. As *SiteTypes* may associate a single domain with multiple categories, it returns a set. For example, *SiteTypes*('*dl.acm.org*') returns {software/technology, educational institution, research/reference}. Popular sites `nytimes.com` and `foxnews.com` are both labeled as {news/media}. Then,

$$Flexibility_{site\_type}(c) = \left\| \bigcup_{s \in SourceSites(c)} SiteTypes(s) \right\|. \tag{12}$$

The minimum value of each granularity of Flexibility for a particular curation product is 0. The maximum is the largest number of sources of granularity *g* achieved among the curations created by participants in a particular study, *C*.

4.1.3. Novelty. Novelty measures the uniqueness of an idea as compared to other ideas. Although this is hard to quantify in the abstract, in the context of a particular study, measuring novelty becomes tractable. The key is a "master list" of all of the answers, with an inverted index showing which participants provided each answer. Thus, we measure Novelty in a manner similar to information retrieval's IDF [Salton and Buckley 1988]. The Novelty metric is globally normalized for each study.

Novelty is measured using three different types of digital found object features, *t*,

$$T_{novelty} = \{image, document, site\}; t \in T_{novelty}, \tag{13}$$

producing the three-element $Novelty_T$ vector, $Novelty_{image,document,site}$. $Novelty_{image}$ measures the average uniqueness of image clippings within a curation. $Novelty_{document}$ measures the average uniqueness of clipping source documents. $Novelty_{site}$ measures the average uniqueness of the Web sites, by domain, from which clippings are curated.

We already have defined sets of features for *source document* locations (Equation 8) and *sites* (Equation 10). We need to form a similar set for *image* locations:

$$ImageUrls(c) = \bigcup_{r_{image} \in c} r_{image}[image\_url]. \tag{14}$$

Then, we define a vector of sets of features for each of these types:

$$F_{images,documents,sites}(c) = \left[ImageUrls(c) \;\; SourceUrls(c) \;\; SourceSites(c)\right]. \tag{15}$$

To compute novelty, for each found object feature for each of the specified types, build inverted indexes that show for each feature of each type the curations containing it:

$$Occurrences_t(url, C) = \{c | c \in C \wedge url \in F_t(c)\}. \tag{16}$$

The magnitude of *Occurrences* measures the usualness of the found object feature. Thus, to compute uniqueness of the feature (i.e., the novelty), take inverse magnitude:

$$FeatureNovelty_t(url, C) = \frac{1}{||Occurrences_t(url, C)||}. \tag{17}$$

The theoretical minimum value of each feature type of Novelty for a curation is $1/n$. The maximum is 1 when a novelty feature appears in only one curation. For example, the more IBI curation products that contain a particular image clipping, the lower its Novelty, $FeatureNovelty_{image}(o_{image}[image\_url], C)$. Likewise, the more curations that contain clippings curated from a particular document or site, the less novel that source.

Then, for each curation, compute $Novelty_t(c, C)$, for each type t, by aggregating the mean Novelty scores among instances of readymade feature type t in that curation:

$$Novelty_t(c, C) = \frac{\sum_{url \in F_t(c)} FeatureNovelty_t(url, C)}{||F_t(c)||}. \tag{18}$$

The result is a triple of normalized values for each feature type, specifying $Novelty_{image,document,site}$. Each value is a rational number on [0, 1].

4.2. Holistic Ideation Metrics of Curation

Holistic metrics are designed to evaluate the creativity manifested in the assemblage of the curation product of an IBI task—that is, how the elements are put together. Each curation is presently evaluated for each holistic metric by independent raters on a 0–3 scale. Each rater is blindly assigned a set of curations, equally distributed across independent variable conditions, without indication of which curations are from which condition. The holistic metrics' reliability—their suitability for producing dependent variables for studies—is determined by raters' consistent performance [Nunnally and Bernstein 1994]. Computing inter-rater reliability measures the consistency of the application of holistic metrics, providing a basis for their validity in any one case, and for their reuse across evaluation contexts.

We develop four procedural guidelines for raters to reliably derive holistic ideation metrics of curation: explicit and clear criteria, mutual independence, the round-down principle, and calibration. To enable effective independent rating, researchers must carefully formulate *explicit and clear criteria* for each point for each metric. The present criteria have been developed iteratively, across multiple experiments and years. To facilitate their application in practice, work to define holistic metrics of curation to be *mutually independent*. Otherwise, raters may experience uncertainty and assign multiple points for the same feature of a curation product, reducing the accuracy of evaluation. The *round-down principle* says, when achievement of a criterion seems partial or incomplete, for consistency, raters do not award the associated point.

Raters of an experiment must convene to *calibrate* assessments. At the start of holistic metrics evaluation, pick a small set (2–4) of initial examples from the dataset. Each rater scores each example. Then, raters then meet to calibrate assessments—that is, to work out consistent contextualized mutual understanding of the criteria—given the IBI task(s) at hand. Iterate calibration meetings as necessary. Forming consensus as such is essential for developing aligned results through subsequent blind ratings.

The present four holistic metrics of curation are Emergence, Relevance, Visual Presentation, and Exposition. Emergence addresses combination of elements leading to new and creative ideas. The other holistic metrics, Relevance, Visual Presentation, and Exposition, function as dimensions of quality in Shah et al.'s parlance. Relevance measures how well curated information matches the topic of the IBI task, and with how much depth. Visual Presentation points address how well the curation works as a visual medium of communication. Exposition measures how well the meaning of what is curated is explained, and why, including how themes are developed.

Exposition is equivalent to Written Presentation. Here, we present explicit and clear criteria for the assessment of each holistic metric.

4.2.1. Emergence. We formulate an Emergence ideation metric of curation, building on Guilford's articulation of the synthesis and reorganization of ideas as an essential aspect of creativity [1950]. Nonemergent combinations only include inherited properties of component concepts [Hampton 1997]. Emergent features are those not present in component concepts, revealed only by combining certain concepts. An example is the gossamer support beam. Support beams, which bear heavy loads, are expected to be massive. The term gossamer support beam emerged in the field of spacecraft design, where design of lightweight structures is imperative. *The Emergence metric of curation measures phenomena in which combinations exhibit characteristics not present in individual elements*, leading to new perspectives, discoveries, and inventions.

Combining ideas has long been of central importance in work on creativity, particularly in brainstorming [Osborn 1963]. For Boden [2003], making unfamiliar combinations is the first form of creativity. Emergence is connected to the "searching for a representation and encoding data in that representation to answer task-specific questions" that Russell et al. [1993] identified as essential to sensemaking. Emergence directly connects to synthesis, addressing creative combination.

We extend the prior method of Kerne et al. [2008] for measuring emergence in information compositions curated to answer ideation tasks. Emergence points are awarded as such:

(1) The first, if a curation develops explicit relationships among digital objects by connecting multiple subsets of heterogeneous elements in groups.
(2) The second, if groups of curated found objects and annotations juxtapose or synthesize concepts. Visual and conceptual juxtapositions both count:
    (a) If the curation connects two concepts together that are otherwise NOT directly connected, this point should be awarded (e.g., intelligent cruise control + GPS navigation IS a synthesis).
    (b) Connections can be visual or conceptual.
(3) The third, if the juxtaposition or synthesis develops new characteristics not previously present in individual digital objects.

4.2.2. Relevance. The Relevance metric indicates how well a curation addresses the specified IBI task. In this research, Relevance may function less as a component of creativity than as a backstop, which detects interjection of random material by participants who did not take a task seriously. The Relevance metric of curation can directly measure significant phenomena for tools that address aspects of work other than ideation, such as search, sensemaking, and analytics. Relevance points are awarded as such:

(1) The first, if the curated found objects are directly related to the IBI task at hand:
    (a) Found objects are intentionally chosen.
    (b) The curation contains few off-topic found objects.
    (c) If the topic of the curation itself is not apparent on looking at the curation, then this point is not awarded.
(2) The second, if the found objects explore connected topics, detailing implications, such as social, economic, cultural, creative, or technical, which contextualize the curation product response to the task question.

(3) The third, if curated found objects and annotations go beyond conveying detailed implications to create an intentional, coherent on-topic focus fulfilling the task.

4.2.3. Visual Presentation. Curation can be presented using visual media of expression and communication. The Visual Presentation criteria are loosely based on design principles of layering and separation articulated by Tufte [1990]. Visual Presentation points are awarded as such:

(1) The first, for the creation of clear spatial relationships using white space.
(2) The second, if a majority of curated elements are positioned, aligned, and arranged to effectively improve visual legibility:
    (a) Arrangement in lines, grids, or other shapes grants this point.
    (b) If spatial relationships are visible without much white space (not enough to give a point for (1)) then this point is granted for arrangement.
    (c) If point (1) is granted, the arrangement must go beyond white space and be somehow consistent in improving visual design to grant this point.
(3) The third, if the curation creates coherent layers and attains legible visual organization through effective use of multiple visual features (e.g., color, size, and font styling):
    (a) No point awarded if there is no layering.
    (b) More than one layer must be present for this point to be awarded.
    (c) If point (1) or (2) are granted, the layering must go beyond the already granted point(s) to improve visual design.
    (d) If points (1) and (2) are not awarded, and only a single visual feature is used to create layers, this point is awarded.

Depending on the curation medium and tool, users may be afforded more or less control over Visual Presentation. The user experience of curation products with regard to Visual Presentation is impacted both by the medium and by how it is used.

4.2.4. Exposition/Written Presentation. Curation combines visual and written communication. To complement Visual Presentation, we instituted the Exposition metric in consideration of expository writing programs in undergraduate education and associated criteria for good written presentation [Harvard College Writing Program 2011]. Exposition measures how well the curation uses text to inform, describe, explain, and expand on constituent ideas. Exposition points are awarded as such:

(1) The first, if the curation explicitly states themes with text:
    (a) Each theme must be supported by more than one found object.
    (b) More than one theme must be stated.
(2) The second, if more than one of the stated themes are further developed and explained through text:
    (a) Text explains themes associated with groups spatially close to them, or to the curation as a whole.
    (b) Explanations should resolve ambiguity regarding themes and/or implications.
    (c) Only consider if the curation was granted a point for (1).
(3) The third, if annotations compare the themes, or describe them, restating them in different words or using literary techniques, such as symbolism, allusion, hyperbole, simile, metaphor, or satire:
    (a) Only consider if the curation was granted a point for (2).

4.3. Statistical Methods for Testing the Significance of Ideation Metrics' Distributions

The IBI evaluation methodology produces experiments that compare curation products made with a new IBI support environment to those created with some pre-existing environment, the control condition. Such experiments investigate, through analysis of curations with elemental and holistic ideation metrics, hypotheses that the new environment better supports IBI. The values of each ideation metric for all curation products result in a distribution for each experimental condition. For holistic ideation metrics, which involve expert raters, also derive a distribution of each rater's assessments for each metric.

We present our approach to statistical analysis of data resulting from derivation of elemental (Section 4.1) and holistic (Section 4.2) metrics of curation. Since the present procedure for deriving holistic metrics utilizes expert raters, this includes measuring inter-rater reliability. We also address issues involving multiple comparisons.

4.3.1. Comparing Ideation Metric Distributions Across Conditions. Instead of using the t-test to parametrically compare means, we prescribe more rigorous Wilcoxon tests to investigate the hypothesis that the distributions are significantly different across conditions for each metric.[2] Wilcoxon is used to derive ranks through arithmetic comparison of pairs of ideation metric values, before comparing their distributions.

For each of the Fluency, Flexibility, and Novelty elemental metrics of curation, we first compare conditions using a Wilcoxon test of the null hypothesis, H_0, that the distributions are identical. If we reject H_0, the difference between conditions is significant. Comparison of means then shows which condition corresponds to a higher value of that metric.

For each of the holistic ideation metrics of Emergence, Relevance, Visual Presentation, and Exposition, there may be multiple values, because the expert raters will not always agree. Here, we use the distributions of the *mean rating* for each curation product as the basis for the Wilcoxon comparisons of the distributions between conditions. The validity of using the means of the ratings depends on their reliability.

4.3.2. Comparing Ideation Metric Distributions Across Raters: Inter-rater Reliability. We must compare the distributions for each curation, across raters, to measure the reliability of their performance, and thus of the holistic metrics. We therefore produce an inter-rater reliability statistic, which measures the consistency of the raters. To derive inter-rater reliability, we must identify a suitable method for computing correlation coefficients for each pair of N raters [Hays 1988]. Cohen's kappa for two raters and Fleiss' kappa for more than two raters [Hays 1988] are not appropriate, because these methods operate on categorical or ordinal data, not interval data, such as the holistic ideation metrics. Spearman rank correlation is also inappropriate, because it assumes that there are no ties among ratings for a particular instance [Hays 1988].

The method of Parzen and Mukhopadhyay [2012] enables computation of correlation coefficients for ranked interval data. To apply this method, transform the sample space by applying a mid-rank operation to the set of ratings for each holistic metric for each curation, across raters. Next, compute correlations for each pair of raters for each metric to measure pairwise agreement. Use these correlations to produce an $N \times N$ correlation matrix for each metric (see example, Table I). This correlation matrix shows

[2]When evaluating IBI support environments, researchers can choose to conduct between-subjects or within-subjects experiments. Different versions of the Wilcoxon test are appropriate for each case [Hays 1988]. For between-subjects experiments, analyze differences in distributions between two independent samples with Wilcoxon rank-sum, also known as Mann-Whitney. For within-subjects experiments, differences between distributions are measured by applying Wilcoxon signed-rank on participants' paired results.

Table I. 4 × 4 Inter-Rater Correlation Matrix Showing Level of Agreement Among Raters $r1$-$r4$ for the Emergence Ideation Metric from Curation Products in the Provocative Stimuli Experiment (Section 7)

	$r1$	$r2$	$r3$	$r4$
$r1$	1.000	0.799	0.900	0.926
$r2$	0.799	1.000	0.903	0.871
$r3$	0.900	0.903	1.000	0.947
$r4$	0.926	0.871	0.947	1.000

Note: The matrix is symmetric around its diagonal, whose values are the identity, since it compares raters to themselves. Rater pair $r1$ and $r2$ is the least correlated, whereas rater pair $r3$ and $r4$ is the most correlated. The overall inter-rater reliability for Emergence was the mean of these values, .891.

how consistently each pair of raters performed for that metric.[3] Then, compute the mean of these correlation matrices over the set of holistic metrics to measure overall agreement for each rater pair. Next, compute the mean of correlation coefficients from rater pairs (those that lie outside of the diagonal). This aggregates the correlation for all raters, forming a measure of overall inter-rater reliability for the experiment.[4]

4.3.3. Multiple Comparisons. When many dependent variables are measured for a single experiment, concerns arise that by chance a statistical test may inappropriately reject the null hypothesis, a false positive (Type I error). Thus, to draw conclusions from the data of a single metric, while computing many, calls for a correction for family-wise error, such as Fisher's protected Least Significant Difference test or his Bonferroni procedure [Hochberg and Tamhane 1987].

In the present research, for each experiment, we discover positive correlations across many metrics. These independent comparisons reinforce our conclusions. Further, some of these possess extremely strong confidence levels, which can be said to protect [Hockberg and Tamhane 1987] the other statistics from requiring more stringent difference criteria.

EPISODE II: CASE STUDY: MIXED-INITIATIVE INFORMATION COMPOSITION

This episode develops a case study investigating ideation metrics of curation in the context of tools using the medium of information composition. As sensitizing concepts, we articulate two creative cognition challenges to IBI—overcoming fixation, the problem of getting stuck in mental ruts, and the need to synthesize findings—to think across results from different sources. In response, Section 6 develops principles of mixed-initiative information composition for supporting creative exploration and curation by integrating automated agents of information retrieval and visualization with tools for curation.

We conducted two experiments to evaluate MI^{2}C, employing the methodology of performing IBI tasks, then evaluating curation products (Section 2.5), using ideation metrics (Section 4). Section 7 presents an experiment addressing the hypothesis that MI^{2}C constitutes provocative stimuli that help people overcome fixation and achieve ideation. Section 8 presents an experiment addressing the hypothesis that MI^{2}C bridges the synthesis gap (Section 5.3), helping people be creative by integrating a curation space for connecting intermediate results with an information visualization.

[3]Expert rating teams can use pairwise correlation coefficients to evaluate how well each rater and pair of raters is performing, in terms of coming to agreement, for each metric, as well as overall. This may be used as the basis for tuning the specification of ratings criteria when raters do not sufficiently agree.

[4]Note: If two distributions are entirely uncorrelated, their covariance and correlation coefficient are zero [Parzen 1960].

5. CREATIVE COGNITION CHALLENGES: SENSITIZING CONCEPTS

We identify two creative cognitive challenges for IBI support environments. One challenge involves overcoming fixation—that is, how people get stuck while trying to ideate. Provocative stimuli constitute a form of relief. The other challenge involves supporting users in traversing the synthesis gap, so they fulfill their needs to generate emergent ideas by combining intermediate concepts and findings.

5.1. Fixation

Fixation means experiencing an obstruction to ideation [Smith and Blankenship 1991]. Fixation does not refer to inability to accomplish a task, or lack of essential knowledge. Rather, fixation is a mental block that prevents someone from completing an otherwise doable job.

A comprehensive experimental program has shown that cognitive fixation involves inappropriate use of knowledge for a task, including conceptual knowledge, memories of events, and learned skills or habits (e.g., Jansson and Smith [1991]; Smith et al. [1993]; Smith and Tindell [1997]). Cognition automates responses for accomplishing habitually encountered tasks, allowing the mind to free up resources [Smith 1996]. Because automatic responses use little or no conscious attention, they may not be subject to conscious intentions. Showing people misleading clues makes it harder to solve rebus puzzles [Smith and Blankenship 1989] and Remote Associates Test problems [Smith and Blankenship 1991], used to assess creative ability (e.g., Mednick [1962]). Making people aware of the source of fixation does not help them think of alternatives. Creative idea generation and design can likewise be constrained by viewing examples [Jansson and Smith 1991; Smith et al. 1993; Kohn and Smith 2009, 2011].

5.2. Relief from Fixation: Incubation and Provocative Stimuli

Relief from fixation is achieved via cognitive restructuring, a shift in the perceptual or conceptual structure of an object or problem. Seeing the problem from a restructured perspective can lead to an unexpected solution. Stimuli encountered outside of one's normal task environment might serendipitously provoke or trigger ideas that had been blocked by fixation. Remedies for fixation include incubation and provocative stimuli. With *incubation*, fixation-breaking stimuli come simply from time away from the task. With *provocative stimuli*, the environment is augmented by encounters with new material designed to help the user overcome fixation and generate new ideas.

The forgetting fixation theory states that incubation allows one to temporarily put out of mind fixated elements of a problem-solving mindset [Kohn and Smith 2009, 2011; Smith and Blankenship 1989, 1991; Smith and Vela 1991; Vul and Pashler 2007]. The opportunistic assimilation theory [Seifert et al. 1995] attributes incubation effects to serendipitously encountered stimuli that serve as clues for solutions. Initial failures leave cognitive representations of an unsolved problem in a partially activated state. When a useful clue is encountered, it triggers memory of the unsolved problem, which is resolved [Seifert et al. 1995; Smith and Blankenship 1989]. However, finding appropriate clues is not easy.

Provocative stimuli are perceived aspects of an environment that help a person overcome fixation, achieve insight [Mednick 1962], and experience ideation. For example, George de Mestral was inspired to invent Velcro when he saw the burrs in his dog's fur after a hike [Velcro 2013]. Archimedes saw water rise in the bath when he got in. He discovered the displacement principle and exclaimed, "Eureka!" [Vitruvius 1914].

C-Sketch is a brainstorming method that mitigates fixation in groups through carefully structured phases of individual and group work [Shah 1998]. A team is given a design problem. To avoid fixation conformity effects, each participant first sketches individually, independently of the others. After a sketch is completed, it is passed to the right. Each participant sketches a second design idea, contextualized by the intermediate creative product received. The process of passing and adding to sketches continues. Shah et al. [2001] used ideation metrics to discover that teammates must sketch independently in this way before collaborating in order for the stimuli of encountered ideas to improve creative ideation.

5.3. The Synthesis Gap: Information Visualization and Visual Analytics

In addition to overcoming fixation, another creative cognition challenge is to support synthesis, which involves combining concepts and intermediate findings to produce emergent ideas (e.g., Wilkenfeld and Ward [2001]). Findings in multiple fields underline the importance of this challenge. Scientific research involves generation, examination, comparison, and analysis of representations to synthesize and express new ideas [Springmeyer et al. 1992]. The "agenda for visual analytics research" of Thomas and Cook [2006] prioritizes the need to help users "discover unexpected and missing relationships." Amar and Stasko [2004] similarly introduced the term *analytic gaps*, which are obstacles faced by information visualization tools, as a result of limited affordances and predetermined representations for working across datasets and domains, to supporting high-level tasks such as decision making. Decision making involves iterative cycles of problem formation, data gathering and analysis, synthesis of potential actions, further analysis, and selection of a particular course of action [Simon 1960].

Information visualization possesses a *synthesis gap*: a need to support comparison, emergence, and ideation. Synthesis is not part of analysis. Synthesis is a complementary cognitive process that involves putting together elements to make up a complex whole [Oxford University Press 2013]. Despite the underlying epistemological clash, the agenda of Thomas and Cook [2006] for visual analytics research prioritizes the need for new "methods to synthesize different types of information from different sources ... so users can focus on the data's meaning in the context of other relevant data."

The present research investigates clipping found objects as a representation for the intermediate findings that people engaged in IBI need to connect and expand upon. Curation of clipping found objects constitutes a medium for synthesis. It affords focusing on data's meaning in the context of other relevant data. We design MI^{2}C to support synthesis by spatially and semantically integrating spaces for curation and visualization, forming an alternative to the usual separate application windows (Section 8). Sandbox [Wright et al. 2006] similarly addresses the synthesis gap, providing integrated space for collecting representations of contextualized intermediate infovis results. However, Sandbox uses arrows to explicitly express relationships among collected elements rather than composition's implicit visual design features, which employ ambiguity to support flexible interpretation [Gaver et al. 2003]. Although implicit and explicit structure both have value, Marshall and Rogers [1992] observed that users engaged in sensemaking tasks avoid formal structure in lieu of spontaneous spatial organization of information objects. They established that only supporting explicit structure imposes limitations on knowledge formation.

5.4. Implicit Structure Visualization

Card et al. [1999] define *information visualization*, quite generally, as 'the use of computer-supported, interactive, visual representations of data to amplify cognition." However, examining their compendium of the field, as well as proceedings of the InfoVis conference, we mostly find a more limited paradigm. In *explicit structure visualization*, which is typified by scatter plots, heat maps, and node-link diagrams, quantities and relationships are directly represented. Explicit structure visualization is powerful, yet incomplete, because it does not leverage implicit structure's ambiguity and flexibility.

In contrast, *implicit structure visualization* presents content directly, expressing relationships through visual elements, including images, spatial relationships, and blending [Webb and Kerne 2011]. Implicit structure visualization algorithms abstractly map parameters from result sets to visual presentation. They give the user a gestalt sense of relationships, without direct conveyance of parameters through labeled axes. They use objects and ambiguous relationships to tell visual stories, supporting flexible interpretations. Implicit structure visualization is useful when people perform ill-defined problems or lack complete a priori cognitive schemas.

Prior research can be interpreted in terms of implicit structure visualization. Design Galleries constituted implicit structure visualization by using multidimensional scaling to project a set of design alternatives onto a 2D space [Marks et al. 1997]. The Bohemian Bookshelf represented a digital book collection through implicit structure visualization, arranging books based on page count into a loosely stacked pile [Thudt et al. 2012]. Relationships between books were implicitly conveyed through spatial positioning, color, and size. Skog et al. [2003] developed ambient information visualizations as " informative art," inspired by paintings of Piet Mondrian. They mapped the number of emails and current temperature to the size and color of rectangles. Their intention was not to precisely represent the data, but to blend it with the surrounding environment, provoking ambiguous, contextualized interpretations.

6. MIXED-INITIATIVE INFORMATION COMPOSITION (MI^{2}C) INTERFACE

We take a mixed-initiative approach in response to the creative cognition challenges of fixation and the synthesis gap. We augment information composition by integrating information retrieval and visualization agents with direct manipulation curation. We develop principles of mixed-initiative information composition, building on the MI^{2}C architecture of Kerne et al. [2008]. A scenario illustrates MI^{2}C experiences. We control conditions for conducting IBI laboratory studies. We use the challenges to IBI to form hypotheses and design experiments for MI^{2}C.

6.1. Principles of Mixed-Initiative Information Composition

To elucidate user experiences of curation augmented by information retrieval and visualization agents, we invoke the "principles for mixed-initiative user interfaces" of Horvitz [1999]. These include value-added automation, timing services based on the user's attention, minimizing the cost of poor guesses, efficient agent-user collaboration, dialogue to refine results, uncertainty about user goals, and working memory of recent actions. We preface these with a principle of our own: give the user space.

6.1.1. Give the User Space. The MI^{2}C user engages directly in information composition in the central curation space. As with Duchamp's creative process, direct manipulation curation begins with choosing readymades (Section 2.5). These digital found objects are Web pages and clippings that represent them, which the user deems meaningful in a task context. The user curates these clippings into the composition via drag and drop. A Firefox plug-in passes the source document URL to the application, which uses meta-metadata [Kerne et al. 2010] to enhance readymade clippings with additional metadata, such as the title and description of the source Web page, and in some cases further details, such as the authors, abstract, and citations of a scholarly article. The user uses interactive tools to arrange, annotate, and blend collected found objects.

6.1.2. Value-Added Automation: Implicit Structure Visualization as Provocative Stimuli. The MI^{2}C interface enables the user to employ automated agents designed to help by stimulating her creativity in the course of IBI tasks. The user initiates the retrieval and visualization agents by specifying seeds. In the present studies, the seeds are searches.

Concurrent to the user's direct curation, the agents retrieve and visualize additional found objects. Using the information extraction semantics provided by the meta-metadata language and architecture [Kerne et al. 2010], the retrieval agents submit queries to Google or Bing and derive semantics from search engine results. The role of meta-metadata is to extract metadata from particular Web sites into structured data models, to form clippings that connect metadata with meaningful images and text passages, and to provide associated term vectors [Salton and Buckley 1988].

Agents use the term vectors and a corresponding model of the user's interests, built in response to interaction with the In-Context Slider (Section 6.1.6), to make decisions about what to collect and how to visualize. The model ranks clipping found objects with dot product similarity. Over time, clippings are fed to an implicit structure visualization (Section 5.4) that serves as an automated agent presenting expected and surprising found objects. The visualization is hypothesized to provide provocative stimuli that help users overcome fixation to be more creative.

6.1.3. Timing Services Based on User Attention. MI^2C changes the timing of information retrieval and visualization. The agents retrieve and visualize information incrementally over time, rather than all at once, like a typical search engine or visualization tool. This differs from rendering time series data. It corresponds to the iterative ways in which people actually collect information [Bates 1989; Marshall and Bly 2005].

At regular intervals (once per second initially), a clipping digital found object is automatically chosen for inclusion in the composition's visualization space. The MI^2C apparatus simply enables the user to turn the visualization agent on and off with a tape recorder–like control. The tape recorder transport panel also provides a speed control, which enables the user to adjust the timing of the visualization stimulus.

6.1.4. Minimizing the Cost of Poor Guesses. Stimuli, including "poor guesses," are presented in periphery, where users can simply ignore them. As the space fills, to reduce clutter, the visualization agent fades out and then removes found objects that the user ignores by gradually reducing saturation and opacity.

6.1.5. Efficient Agent–User Collaboration. The MI^2C interface enables efficient agent–user collaboration by employing consistent affordances, the same information composition medium, and coupled spaces for the user's curation and the agent's visualization.

The areas for direct-manipulation curation and mixed-initiative visualization are spatially and semantically continuous, with the user's work in the center of attention and the agent's on the periphery. The curation space, where the user collects information of importance to her and engages in synthesis, is placed centrally to garner focus in creative engagement. Layout of the visualization space at the periphery makes it easy for the user to ignore it when she wants to. The user can simply drag found objects from the visualization to the curation space.

6.1.6. Dialogue to Refine Results. The MI^2C apparatus provides a fluid interface component to enable the user to engage in dialogue about which visualized found objects are of interest. The In-Context Slider enables the user to adjust the level of interest in a clipping, metadata field, or word(s), from -5 to $+5$ [Webb and Kerne 2008]. This brings feature vectors and weights into the user model, tuning ranking operations on found digital objects that feed the implicit structure visualization.

6.1.7. Uncertainty About User Goals. Hortivz [1999] sees uncertainty in the agents' model of the user's goals as a problem. The context of IBI faces a fundamental issue: the user's goals may be indefinite and subject to spontaneous changes, and so uncertain to the user

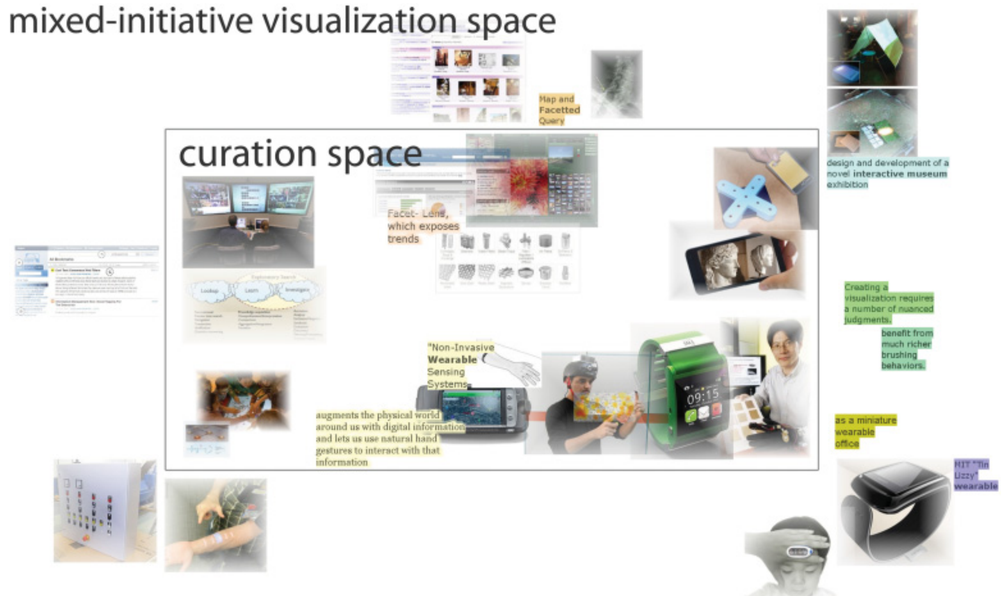

Fig. 3. Kate curates in the medium of information composition. New clippings in the mixed-initiative visualization space on the periphery are automatically placed near related groups in the central curation space.

herself (Section 2.6). Spontaneity and serendipity are essential to processes of creativity and learning, including the formation of emergent combinations (Section 4.2.1). Thus, in IBI, user uncertainty about goals is inherent and beneficial.

6.1.8. Working Memory of Recent Actions: Dyadic Undo. Horvitz [1999] prescribes maintaining a memory of recent actions involving users and provides mechanisms for efficient manipulation of "shared" experiences. Part of minimizing the cost of poor guesses is to make all actions reversible. The MI^{2}C apparatus accomplishes this through a "Dyadic Undo" mechanism that makes the composite of actions by user and agents reversible.

6.2. Scenario

We illustrate the MI^{2}C interface through a scenario. Kate, a computer science Ph.D. student, is researching prior work as a springboard for formulating dissertation ideas. She wants to synthesize HCI, IR, and infovis to transform museum experiences to be more educational, engaging, exploratory, and exciting.

Kate is interested in the roles that wearable computing can play. She searches the Web for wearable computing devices. She browses the first result, on SixthSense [Mistry 2012], a project that augments the physical world with information and enables interaction with hand gestures. Kate sees an interesting image of a man wearing a helmet with electronic equipment. She curates this image clipping found object to represent SixthSense, dragging it from her Web browser and dropping into the curation space (Figure 3, bottom center of curation space). She also curates text from the same page.

Kate curates more wearable computing found objects. She searches for faceted information visualization. She finds Flamenco, a Web-based interface for browsing large collections of items such as documents or photographs [Yee et al. 2003]. She imagines that museum visitors will want the ability to search for additional information on an exhibit, and to find other related exhibits. She curates an image clipping of Flamenco.

Kate continues her exploratory search, browsing and reading results, and curating found objects. As time passes, she begins to fixate. She is having a hard time finding new information that expands the conceptual space of her synthesis. Kate decides to engage the software agents of MI^2C to help her curate more diverse information.

Kate seeds the agents with a set of search queries that cover topics of interest: natural user interface technologies, faceted information visualization, exploratory search, wearable computing devices, and interactive museum.

The agents use her seeds to perform searches and find documents. Images and text are clipped from the retrieved documents and incrementally visualized as digital found objects in the periphery of the composition. Clippings appear from each seed, enabling Kate to explore varied topics and develop a multidisciplinary view.

Suddenly, in the visualization space, Kate sees an interesting diagram of exploratory search [Marchionini 2006]. She drags the clipping from the visualization into her curation. She browses the article, which explains the kinds of human processes involved in exploratory search. The clipping as found object serves as a stimulus, provoking Kate to reflect across topics: natural user interfaces and exploratory search. She notices an image clipping depicting a map with information overlaid. Upon dragging into the curation space, she discovers a conceptual juxtaposition by playfully placing the map image in between the SixthSense demonstrator's fingers. This sparks an emergent idea: museum visitors can interact with maps to discover what is exhibited on each floor, how much time each exhibit takes, and how busy it is.

As Kate curates information, she organizes image and text clippings into groups, creating messy piles of found objects. She adds labels for wearable computing, faceted information visualization, and exploratory search to emphasize the grouping. Kate sees clippings appearing in the visualization space near related groups in the curation space, growing the conceptual space (Flexibility) of the curation.

In the visualization space, near her wearable computing group, Kate notices a text clipping, "non-invasive wearable sensing systems." She brushes her mouse cursor over "non-invasive," activating the In-Context Slider. She expresses positive interest. New clippings appear nearby, depicting less invasive technologies (Figure 3, bottom right).

Kate continues collecting, grouping, juxtaposing, and designing as she develops her synthesis. She adds an annotation to serve as a title: "a new museum experience: wear, explore, visualize, investigate, discover, and learn." Kate saves her composition and emails it to her advisor and colleagues to get feedback on her ideas (Figure 4).

Without MI^2C, Kate would explore each search separately. As in the beginning of this scenario, she would explore results and curate clippings one topic at a time. She would be likely to fixate on topics in her initial queries, which could result in never performing searches with the latter queries. She would probably only browse a small number of results from each query, instead of being stimulated by the first 30 or more. Kate's limited human attention would have to perform all of the work of integrating and exploring diverse topics, without support from MI^2C components.

6.3. Controlled Experimental Conditions: Fixed Searches

Controlling study conditions across participants supports empirical investigation by increasing the chance to observe significant differences. For example, Woodruff et al. [2001] developed an experiment that evaluated "enhanced thumbnails" as an elemental medium for information resources in sensemaking tasks. Their experimental procedure fixed the search queries that study participants explored and cached search results. Dynamic changes in search engine rankings would otherwise lead to differing search results on the same queries, confounding comparison across conditions. They

Fig. 4. Kate's information composition is a curation product of her research topic formulation information-based ideation task, which synthesizes prior work for her dissertation research. She creates messy groupings of digital found objects, juxtaposing different aspects of the research, and providing a holistic view. Each clipping found object includes attribution to its source. She can share the composition-as-ideation-curation with colleagues and her advisor.

Table II. MI^{2}C as Provocative Stimuli Study Conditions:
Information Composition Apparatus for Ideation Support

Time	Provocative Stimuli	Control
10 min	direct-manipulation-only curation with searches	direct-manipulation-only curation with searches
20 min	mixed-initiative information composition integrating curation with retrieval and visualization agents	
no limit	direct-manipulation-only curation: organize, design, & annotate	

considered the user experience of their study as representative of "Web activities [that] significantly impact . . . decisions and actions."

The present studies (Sections 7 and 8) of MI^{2}C likewise controlled conditions by fixing and caching the searches participants explored. Although this limits the overall creativity of experiences of individuals in studies, this limitation, which would not be present in a real-world setting, is fairly distributed across experimental conditions:

—For the Stem Cell Research IBI task (Section 3.2.1), the searches we used were (1) stem cell debate, (2) stem cell research, (3) miotic cell genetics cure embryonic, (4) cloned stem cell genes combat disease, (5) genetics ethics, and (6) regenerative cell DNA sequence.
—For the Liberty and Security IBI task (Section 3.2.2), the searches were (1) big brother government news, (2) liberty versus security, (3) homeland security, (4) freedoms lost to terrorism, (5) mass surveillance, and (6) civil liberties.

6.4. IBI Experiments to Evaluate MI^{2}C

We develop two laboratory experiments, each addressing the contribution of components of MI^{2}C to support for overcoming challenges to IBI. One study investigates the *hypothesis that MI^{2}C serves as a source of provocative stimuli* to help users overcome fixation and be more creative in performance of IBI tasks. The second study investigates the *hypothesis that MI^{2}C bridges the synthesis gap by semantically and spatially integrating visualization and curation*. In all conditions across these studies, participants create information compositions as curation products of IBI tasks.

7. EXPERIMENT: MI^{2}C AS PROVOCATIVE STIMULI

7.1. Experimental Design

We used Shah et al.'s C-Sketch (see Section 5.1) as a paradigm to develop a provocative stimuli investigation of the MI^{2}C apparatus (Section 6.1). To maximize creative output, Shah et al. [2001] had designers work individually, before exposing them to each other's work. Kohn and Smith [2011] developed consistent findings: group brainstorming tends to lead to fixation if participants don't first work individually.

The problems with group brainstorming gave us concern about software agents working with a human on an IBI task. As with human collaborators, if the agents start, the user could become fixated, short-circuiting her own process of inception.

Thus, we developed a provocative stimuli experiment design to investigate the impact of MI^{2}C as a remedy for fixation. The independent variable that we manipulated was information composition apparatus for ideation support: Provocative Stimuli versus the Control, direct-manipulation-only (Table II).

In the *Provocative Stimuli condition*, in Phase 1, participants perform direct-manipulation-only curation, without the help of agents (10 minutes). Next, they engage in MI^{2}C, in collaboration with information retrieval and visualization agents (20 minutes). The whole mixed-initiative information composition interface (Section 6)

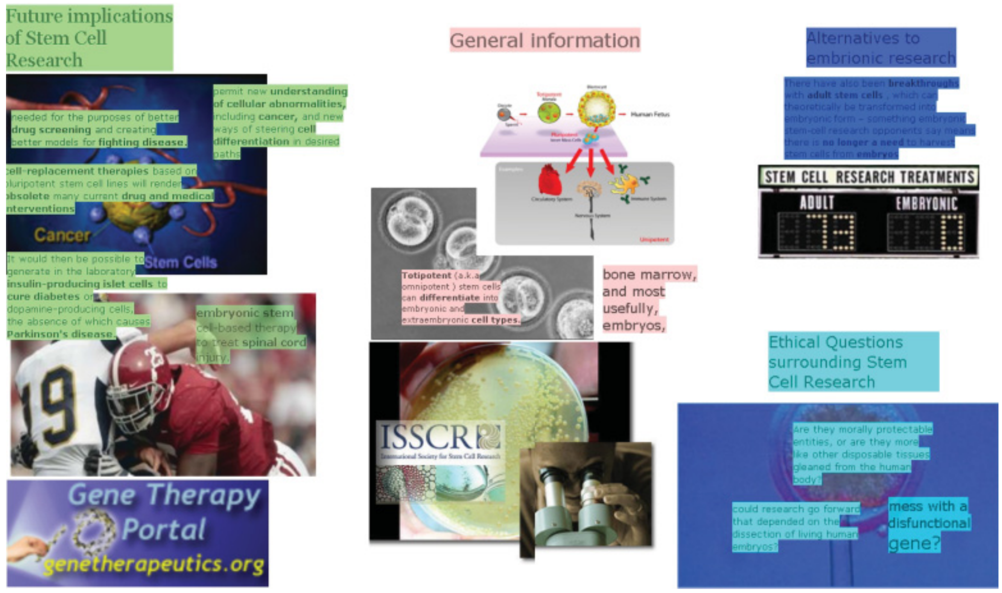

Fig. 5. Information composition curation product for the Stem Cell Research IBI task, Provocative Stimuli condition (Section 7.1). Elemental metrics are interpreted using the means of this experiment (Section 7.3). $Fluency_{image,text}$ of [10 11] is near the means. $Flexibility_{document,site,site\_type}$ of [14 12 5] also represents near mean diversity of curated information sources. $Novelty_{image,document,site}$ is [.32 .12 .08], indicating curation from commonly used documents and web sites, with some unique image clippings, such as one of American football. Holistic metrics, on a 0–3 scale, measure [2 3 3 2] for Presentation, Relevance, Emergence, and Exposition, respectively. Presentation points are given for use of white space to convey relationships and layering through use of color and font size. An emergent idea connects treatment of sports injuries to implications for stem cell research.

is at hand. The implicit structure visualization is presented in the peripheral visualization space (Figure 3). The participant continues to work in the curation space, in the center, refining her answer(s) to the IBI task. In the final phase, whose duration is open ended, the participant again engages in direct-manipulation-only curation. To refine the curation product, she is instructed to finish creating her composition, organize elements, design the composition, and annotate to explain and motivate the answer.

The *Control condition* is similarly structured, except that it omits MI^{2}C, leaving two phases. During Phase 1, which lasts for the combined duration of Phase 1 and Phase 2 of the Provocative Stimuli condition (30 minutes), the participant engages in direct-manipulation-only curation. The open-ended final phase is the same as before. The Control condition is a subset of the provocative stimuli condition, which interjects MI^{2}C in the midst of the user's information composition curation experience.

During Phase 1 of each condition, the fixed searches (Section 6.3) of the IBI task at hand (Section 3.2) were presented from an initial Web page, which led to typical Google style formatted search results. During Phase 2 of the Provocative Stimuli condition, the information retrieval agents were fed with the same search query seeds and cached search result Web pages as in Phase 1, except that the Web pages the user had already browsed were omitted from the agents' collection of found objects and so from the visualization. During Phase 3, the last phase of both conditions, search results were not accessible, but participants, engaging in direct-manipulation-only curation, could still organize, design, and annotate the curation product.

Table III. Elemental Ideation Metrics of Curation for Creative Products in MI^{2}C as Provocative Stimuli Experiment with Wilcoxon Rank-Sum Statistics

Metric	Provocative μ	SE	Control μ	SE	W	$p <$
Image Fluency	14.9	2.0	5.5	0.6	447	.00001
Text Fluency	14.7	1.0	13.6	1.1	294	.35
Document Flexibility	19.5	1.8	5.5	0.5	492	.00001
Site Flexibility	16.8	1.6	5.0	0.5	488	.00001
Site Type Flexibility	7.7	0.5	4.9	0.4	414	.00021
Image Novelty	.28	.03	.17	.03	362	.012
Document Novelty	.27	.02	.16	.03	382	.0033
Site Novelty	.16	.02	.08	.02	400	.00076

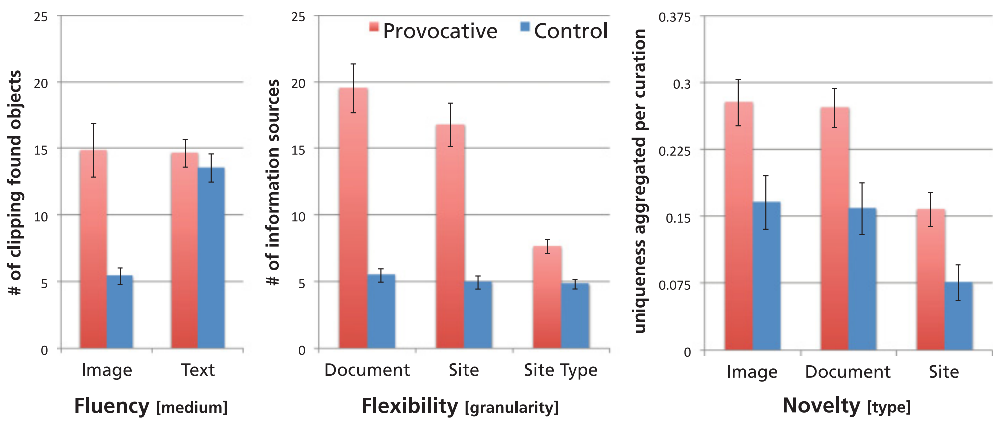

Fig. 6. Mean elemental ideation metrics of curation for creative products in MI^{2}C as provocative stimuli experiment. Curations from the Proactive Stimuli condition (in red) possess significantly greater IBI than the Control condition (in blue) for all elemental metrics, except for Text Fluency. Error bars show standard errors of the mean.

7.2. Procedure

We conducted a between-subjects experiment with 44 participants, according to the IBI support environments procedure (Section 3.3). The participants were undergraduate introductory psychology students who gained course credit by choosing to participate. They were not previously familiar with the research or the researchers.

Each participant performed the Stem Cell Research IBI task (Section 3.2), creating an information composition as curation product (see example, Figure 5). Half the participants were subject to the Provocative Stimuli condition MI^{2}C apparatus for IBI support. The other half were subject to the Control condition.

Phase 3 had no set time limit. Participants spent an average of 14.2 minutes to finally organize, design, and annotate. There were no significant differences between conditions for Phase 3 duration.

7.3. Results: Elemental Ideation Metrics of Curation

Elemental metrics of curation were computed according to the methods of Section 4.1. Figure 6 graphs the results. Measures of statistical significance in the difference between distributions for the two conditions were calculated using the Wilcoxon rank-sum test. Table III presents means for elemental metrics and statistical significances of differences. Participants in the Provocative Stimuli condition exhibited greater Fluency, Flexibility, and Novelty than those in the Control condition. All results are statistically significant except for Text Fluency.

Table IV. Holistic Metrics of Curation Measure Ideation in MI^{2}C as Provocative Stimuli Experiment with Wilcoxon Rank-Sum Statistics

Metric	Provocative μ	SE	Control μ	SE	W	$p <$
Emergence	1.71	0.16	1.14	0.19	334	0.025
Relevance	2.77	0.10	2.86	0.07	219	0.481
Visual Presentation	0.99	0.14	1.53	0.18	144	0.020
Exposition	1.38	0.15	1.36	0.13	243	0.990

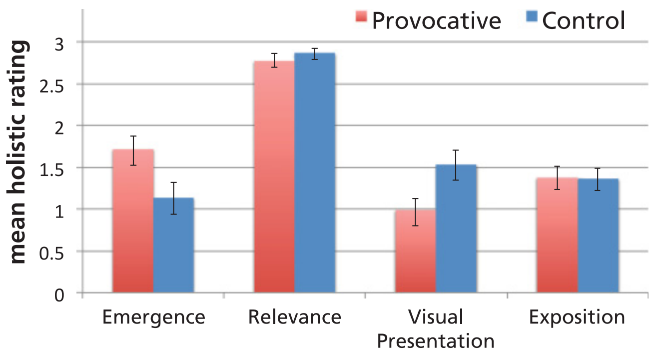

Fig. 7. Mean holistic ideation metrics of curation for creative products in MI^{2}C as a provocative stimuli experiment, with standard error bars. Participants' curations in the Provocative Stimuli condition exhibited significantly increased Emergence (the metric associated with synthesis) and significantly decreased Visual Presentation.

For global interpretation of the elemental metrics for any such curation, we present experiment-wide mean value vectors, which show the average values of each elemental metric across both interface and study task conditions. For example, this enables us, for Figure 5's example information composition, to interpret the $Novelty_{image}$ score of .32 as above average, indicating the presence of some unique images:

$$\begin{aligned}
\overline{Fluency}_{image,text} &= \begin{bmatrix} 10.4 & 14.1 \end{bmatrix} \\
\overline{Flexibility}_{document,site,site_type} &= \begin{bmatrix} 13.0 & 11.3 & 6.4 \end{bmatrix} \\
\overline{Novelty}_{image,document,site} &= \begin{bmatrix} .23 & .22 & .12 \end{bmatrix}.
\end{aligned}$$

7.4. Results: Holistic Ideation Metrics of Curation

Holistic ideation metrics of curation were assessed by four independent, reliably consistent expert raters according to the methods of Section 4.2. Each composition as curation was rated by all four raters. The raters were blind to the information composition apparatus experimental condition used to create each curation product. The mean of each metric was used as the input to statistical tests.

Figure 7 shows the mean Visual Presentation, Relevance, Emergence, and Exposition for all curation products and raters across experimental conditions. Table IV presents holistic metric means and the statistical significance of differences between distributions. The Wilcoxon rank-sum test was used to assess statistical significance. Emergence was found to positively correlate with the Provocative Stimuli condition, whereas Visual Presentation was found to negatively correlate.

Following Section 4.3's methodology for measuring inter-rater reliability, for each holistic metric, we applied mid rank to the ideation metrics assessed for each curation, across raters. We then applied Pearson's correlation over the 44 curated products for each metric, producing a 4×4 correlation matrix. We calculate a mean overall

Table V. Provocative Stimuli Experiment Holistic Metrics Inter-Rater Correlation Matrix

	$r1$	$r2$	$r3$	$r4$
$r1$	1.000	0.809	0.859	0.887
$r2$	0.809	1.000	0.825	0.839
$r3$	0.859	0.825	1.000	0.884
$r4$	0.887	0.839	0.884	1.000

correlation matrix (Table V) from these individual holistic ideation metric rater correlation matrices. Taking the mean of nondiagonal elements, we discovered an average inter-rater reliability correlation of .85, indicating strong rater agreement.

8. MI^{2}C EXPERIMENT: BRIDGING THE SYNTHESIS GAP

We derived the concept of *synthesis gap* (Section 5.3). We pointed out that analysis inherently involves understanding what is, whereas synthesis involves ideation. We observed that the gaps that Amar and Stasko [2004] identified in infovis, such as the worldview gap, are more than "analytic" gaps. We likewise observed that key findings by Thomas and Cook [2006] go beyond analytics, including "[an] urgent [need] to develop a data synthesis capability so [users] can concentrate on the data's meaning," and to support users in "discover[ing] unexpected and missing relationships that might lead to important insights." Synthesis is an essential process in curation and in science.

8.1. Experiment Design

We designed an experiment to investigate the hypothesis that MI^{2}C bridges the synthesis gap and promotes IBI by spatially and semantically integrating implicit structure visualization with curation. The curation space serves as a site of conceptual synthesis, where the user collects and connects intermediate results, and forms emergent ideas. Semantic integration is achieved by enhancing visual clippings with metadata, including links, enabling reflection on multiple intermediate results and return to intermediate contexts, in conjunction with holistic visual presentation of the curation.

Most prior information visualization tools do not explicitly support the user in curating intermediate results, with context, enabling return and reflection. Instead, the user is left to perform curation in a spatially and semantically separate application.

The independent variable was the relationship between the direct curation and automatic visualization spaces of the IBI support tool in the study apparatus: Integrated versus Control (separated) (Figure 8). Across conditions, the spaces allocated for curation and visualization were of identical size. For controlling the visualization, the same tape recorder transport and In-Context Slider mechanisms operated the same way.

In the Integrated condition (Figure 8(a)), there is a single, full-screen window (Figure 8), as per the MI^{2}C interactive system design (Section 6.1). Curation is performed in the center space. Implicit structure visualization runs in the peripheral area.

In the separated Control condition (Figure 8(b)), the implicit structure visualization was presented in a separate window from the curation space. Participants could drag visual clippings from the visualization space and drop in the curation space. However, metadata was not transmitted. This control condition exemplifies typical prior conditions with regard to the separation of information visualization tools from other applications, like MS Word or Powerpoint, which people use for curation and synthesis of intermediate results. If the infovis tool does not support drag and drop is not supported, it would also be necessary to take and crop screenshots.

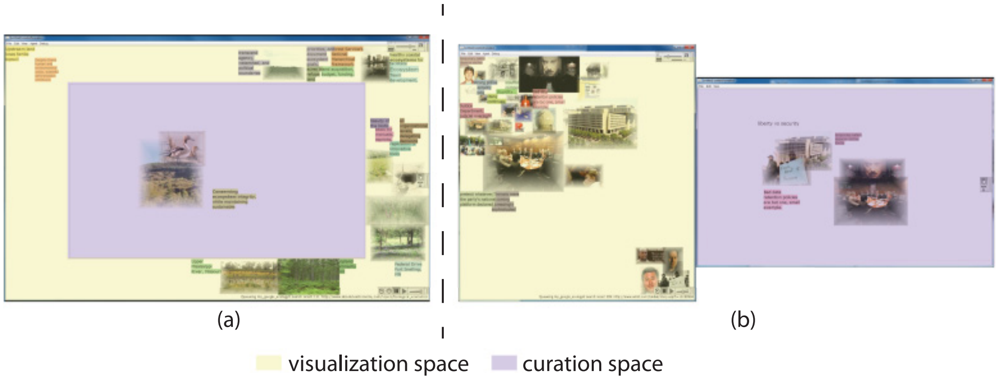

Fig. 8. Bridging the synthesis gap experiment: comparison of spaces in the Integrated and Control (separated) conditions. In (a), the Integrated condition, the direct manipulation curation space is in the center of a single window, coupled to the visualization space on the periphery. In (b), the Control condition, the direct manipulation and visualization spaces are in separate windows. In both conditions, clippings are dragged and dropped from the mixed-initiative visualization space into the direct curation space.

Table VI. Synthesis Gap Study Conditions: Curation and Exploration Apparatus for IBI Support

Time	Integrated	Control
18 min	mixed-initiative information composition with agents	information composition and agents
10 min	direct-manipulation-only information composition: organize, design, & annotate	

8.2. Procedure

We conducted a 2×2 within-subjects experiment across IBI tasks and IBI support apparatus, which were counterbalanced. We followed the IBI support environments procedure (Section 3.3). Each participant performed the Stem Cell Research and Liberty versus Security IBI tasks (Section 3.2). Each participant curated her answer to each IBI task in the medium of information composition (see example, Figure 1).

There were 49 participants in the experiment. They were undergraduate introductory psychology students who gained course credit by participating. The participants were not previously familiar with the research or researchers.

Each experimental condition was divided into the same two phases (Table VI). In the first phase, implicit structure visualization presented new digital found objects. Participants spent 18 minutes interacting with the visualization, curating information from it, and developing ideas. Curation was accomplished by drag and drop from the visualization space to the curation space.

In the second phase of each experimental condition, the visualization was stopped. Participants could still drag and drop across spaces, but no new information appeared. They spent 10 minutes organizing, designing, and annotating to complete the curation. The experiment concluded with postquestions that asked participants to express preferences between the integrated and control MI^{2}C curation tool experiences.

8.3. Results: Elemental IBI Metrics

Elemental ideation metrics of curation were computed according to the methods of Section 4.1. Figure 9 graphs results. Measures of statistical significance of difference between distributions for the two conditions were calculated using the paired Wilcoxon signed-rank test. Table VII presents means for elemental metrics and statistical significances of differences. Fluency, Flexibility, and Novelty elemental metrics demonstrate greater ideation in the Integrated condition. All results are statistically significant except for Site Novelty (which is close; $p < .072$).

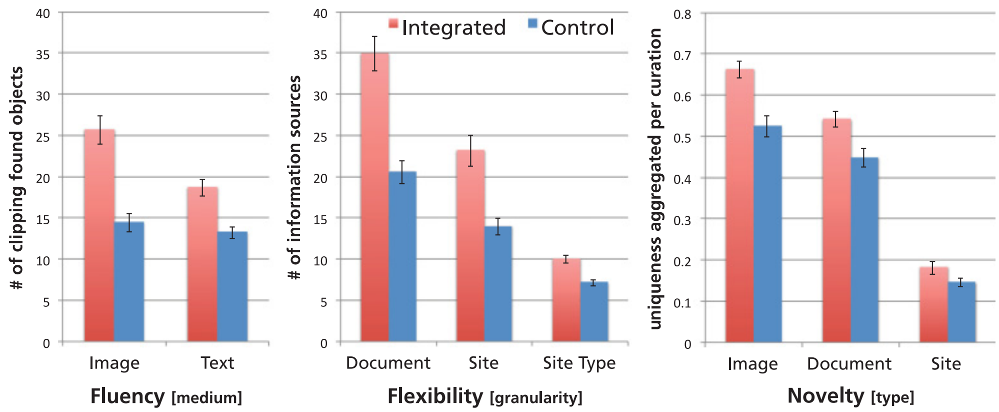

Fig. 9. Elemental ideation metrics of curation for creative products in the synthesis gap experiment: means and standard error bars. We compare the Integrated MI^2C condition (in red) with the separated Control condition (in blue). Curation products from the Integrated condition show significant improvements in creativity.

Table VII. Elemental Ideation Metrics of Curation in the Synthesis Gap Experiment, with Paired Wilcoxon Signed-Rank Statistics

Metric	Integrated μ	SE	Control μ	SE	V	$p <$
Image Fluency	25.7	1.7	14.5	1.1	1,123	.00001
Text Fluency	18.8	1.0	13.3	0.7	1,023	.00001
Document Flexibility	35.0	2.1	20.6	1.4	1,158	.00001
Site Flexibility	23.2	1.9	14.0	1.0	1,037	.00001
Site Type Flexibility	10.1	0.5	7.2	0.4	927	.00001
Image Novelty	.66	.02	.53	.03	1,022	.00001
Document Novelty	.54	.02	.45	.02	928	.00050
Site Novelty	.18	.02	.15	.01	764	.072

Figure 1 depicts an example curation from this experiment. For interpretation of any single curation's elemental metrics, we present experiment-wide mean value vectors:

$$\begin{aligned}
\overline{Fluency}_{image,text} &= \begin{bmatrix}20.1 & 16.0\end{bmatrix},\\
\overline{Flexibility}_{document,site,site_type} &= \begin{bmatrix}27.8 & 18.6 & 8.6\end{bmatrix},\\
\overline{Novelty}_{image,document,site} &= \begin{bmatrix}.59 & .49 & .16\end{bmatrix}.
\end{aligned}$$

8.4. Results: Holistic IBI Metrics

To obtain holistic ideation metrics of curation, each of the 98 curation products from the 49 participants in this within-subjects experiments were rated independently and blindly by two raters, according to the methods of Section 4.2. None of the holistic metrics demonstrated statistically significant differences between conditions (Table VIII).

Again, following the present methodology (Section 4.3) for measuring inter-rater reliability, we applied mid rank and used Pearson's correlation to calculate a 2×2 correlation matrix for each Holistic metric. We calculated the mean of these four matrices to derive an overall inter-rater reliability of .83, indicating a high level of consistency.

Table VIII. Holistic Ideation Metrics for Curation Products in the Synthesis Gap Experiment

Metric	Integrated μ	SE	Control μ	SE	W	$p <$
Emergence	1.47	0.11	1.41	0.12	268	0.691
Relevance	2.13	0.05	2.18	0.06	82	0.604
Visual Presentation	0.78	0.07	0.78	0.09	73	0.811
Exposition	0.62	0.10	0.71	0.10	113	0.459

Note: Paired Wilcoxon tests show no significant differences between conditions.

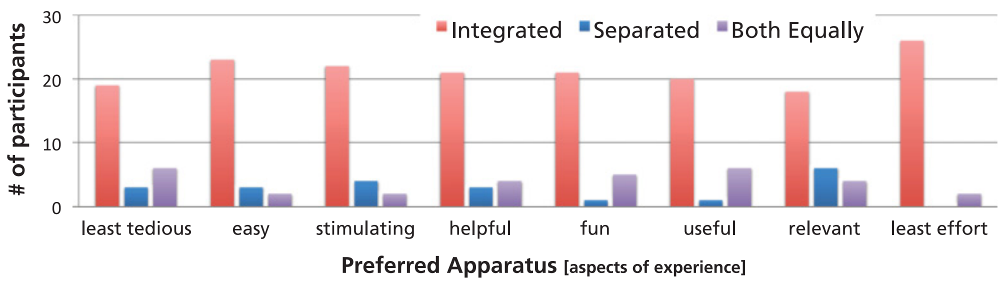

Fig. 10. Experience report data for the Synthesis Gap experiment. Participants compared their experiences in the Integrated MI^{2}C (red) conditions and separated Control (blue), expressing for each of eight aspects of the MI^{2}C experience if they preferred the Control apparatus, the Integrated apparatus, or both were equal. Statistically, participants favored the Integrated condition in all assessed dimensions of experience.

8.5. Results: Experience Reports

We gathered experience report data from 28 participants. Errors in the study apparatus prevented storing responses from the other 21 participants, but this did not interfere with obtaining significant results (Figure 10). We asked participants to express a preference between the Integrated and Control conditions with regard to eight aspects of their experiences. Responses were highly in favor of the Integrated spaces condition, correlating with the elemental ideation metrics. The chi-squared test was used to validate that observed quantitative results from subjective user experience Likert responses are significantly different from the expected results of an equal selection for each response.

Participants experienced the Integrated condition as less tedious ($\chi^2 = 15.5$, $p = 0.0005$), easier to use ($\chi^2 = 30$, $p < 0.0001$), more stimulating ($\chi^2 = 26$, $p < 0.0001$), more helpful ($\chi^2 = 21.9$, $p < 0.0001$), more fun ($\chi^2 = 38.9$, $p < 0.0001$), more useful ($\chi^2 = 6.5$, $p = 0.00388$), easier to collect relevant information with ($\chi^2 = 12.3$, $p = 0.002$), and requiring less effort ($\chi^2 = 44.85$, $p < 0.0001$) than the Control condition.

9. DISCUSSION

The final episode of this article (Sections 9–11) analyzes and synthesizes the design process, methodologies, and results of this work to articulate contributions in creative cognition and HCI research. We discuss findings and develop implications for design. We finish by drawing conclusions and extrapolations for future work.

We consider the creativity support environments evaluation methodology of IBI tasks coupled with metrics of curation products. We connect the results of the present experiments, contextualized by prior experiments, using this IBI methodology, to validate mixed-initiative information composition. We establish hypotheses about how MI^{2}C supports creativity. We also address shortcomings.

9.1. Toolbox of Ideation Metrics of Curation

We initiated a toolbox of ideation metrics for assessing distinct components of creativity in products that people curate in performance of IBI tasks. We classified metrics as elemental or holistic. The elemental metrics, which assess creativity within the found objects that people collect, are derived from prior work by Shah et al. in the context of engineering design. The sources of curated information serve as the basis of new computational methods. We invented the holistic metrics specifically to assess creativity in how elements are put together, addressing the assemblage of a curation. We defined a reliable method for assessment of holistic metrics.

The present toolbox of ideation metrics of curation represents a milestone on an ongoing path. We expect definitions and methods for deriving ideation metrics to continue to evolve. We invite members of the CHI community to join this ongoing endeavor.

9.2. An Empirical Science of IBI Support Environments

A primary finding is that the ideation metrics of curation function as effective discriminants for measuring the efficacy of IBI support environments in different contexts. In addition to the present two studies, the elemental metrics plus Emergence were measured in a prior laboratory study of information composition as a medium for exploration and curation [Kerne et al. 2007], with significant results derived for Emergence and Flexibility. More recently, a field study of a different information composition authoring tool, InfoComposer, invoked the elemental metrics to compare coupled changes in pedagogy and IBI support environment across semesters, deriving significant results for two measures of Fluency, two measures of Flexibility, and three measures of Novelty [Webb et al. 2013].

Thus, IBI methodology has already led to the discovery of significant results in four studies: one field study and three laboratory experiments. Three of the studies measured significant differences in Fluency and Novelty (the fourth did not measure these at all.). Two discovered increased Emergence: information composition as a medium for curation, and MI^2C as provocative stimuli.

All four studies discovered significant changes in Flexibility metrics of curation. Flexibility is increased by using information composition as the medium of exploration and curation, by using visual clippings with contextual metadata as the elements of curation, by augmenting information composition with MI^2C to provide provocative stimuli, and by integrating visualization with curation rather than leaving them separate.

The present methodology of metrics of curation for IBI tasks descends directly from Guilford [1950, 1956, 1968]. It builds on seminal work in engineering design. Shah et al. [2001] used Variety and Novelty to validate the C-Sketch method. Shah et al. [2002] subsequently articulated a space of metrics for design ideation. Shah et al. [2003] and Nelson et al. [2009] continued to refine and apply the design ideation metrics.

The successful application of ideation metrics across information-based and design ideation supports the conclusion that this empirical creative cognition approach can be generalized to diverse ideation support contexts. Given that *science* refers to knowledge established through systematic procedures [Oxford University Press 2013], we have demonstrated that the methodology of evaluation via metrics of users' curation products constitutes a basis for a science of IBI support environments.

9.3. Validation of Hypotheses: Mixed-Initiative Information Composition Supports IBI

We found that with regard to providing provocative stimuli that help users overcome fixation, and bridging the synthesis gap by integrating curation with authoring, MI^2C supports IBI. These hypotheses are validated across experiments by the consistent impact on two measures of Fluency, three measures of Flexibility, and three measures of Novelty.

Novelty and Flexibility metrics across the present experiments show that IBI performed by Provocative Stimuli participants was measurably more diverse and original, and thus more creative. Results of improvement in multiple Novelty metrics show that the provocative stimuli resulting from MI^{2}C's visualization agents did not cause participants to collect the same information, following the agents in a rote manner. Integration of curation and visualization spaces likewise increased Novelty. Thus, MI^{2}C stimulates users to curate different clippings from different Web pages and sites than each other, indicating mutually distinct thought processes and ideas. By comparison, Novelty metrics show that users in the Control conditions fixated on the same clippings as each other. The improvement in multiple Flexibility metrics similarly shows that the provocative stimuli of MI^{2}C and the integration of curation and visualization causes users to engage with more diverse information rather than fixating on a limited space of possibilities.

9.3.1. Provocative Stimuli Hypothesis Validation. The MI^{2}C as a provocative stimuli experiment contributes significant findings regarding Emergence and Novelty, in addition to Fluency and Flexibility. Since Emergence and Novelty directly measure the creation of new ideas, this experiment further demonstrates the power of MI^{2}C as a support environment for creativity on IBI tasks. The Fluency and Flexibility results connect creative process with products. Qualitative and quantitative results of the prior, omnibus field studies add ecological validity to these componentized analytic experimental findings. The provocative stimuli of MI^{2}C were found directly to spur participants to create new ideas, to think outside the box of the provided information.

Previously, Kerne et al. [2008] found that using information composition instead of text lists increased Emergence on IBI tasks. This study replicates validation [Wilson et al. 2013] of the role of information composition as a medium of exploration and curation in promoting Emergence, and extends. Agents automatically collecting and visualizing information in the periphery further increase Emergence.

Cognitive psychology luminary Arthur Glenberg showed that supporting text with homogeneous images promotes mental model formation [Glenberg and Langston 1992]. The more heterogeneous image and text clipping digital found objects in MI^{2}C have a similar effect.

The increase of Emergence in the first experiment, but its absence on the second, validates the Provocative Stimuli experience structure. First, engage the user in self-directed curation. Then, to avert fixation, provide the stimulus of visualization agents in MI^{2}C. This will better support Emergence, the synthesis of new ideas derived by combining existing information.

9.3.2. Synthesis Gap Hypothesis Validation. The synthesis gap experiment results show that MI^{2}C's seamless integration of visualization stimuli with curation increases creativity, indicated by the elemental metrics of Fluency, Flexibility, and Novelty. Integrating curation with visualization caused no ill-effects on any holistic measures. There is no measurable distraction in this integration, as compared to the distraction inherent in switching contexts. Snap together views facilitate connection of datasets and visualizations thereof [North and Shneiderman 2000]. When curation is tightly coupled with visualization, it too "snaps together," not as a single view but as a place to connect heterogeneous items from other views.

9.4. Information Composition Supports IBI

The present studies add to our knowledge of how information composition, as a medium for exploration and curation, supports IBI. We found that MI^{2}C specifically provides provocative stimuli and bridges the synthesis gap. A prior laboratory study showed that as a medium for exploration and curation, information composition promotes IBI through increased Flexibility[5] and Emergence [Kerne et al. 2007, 2008]. A prior field study in a large course on entrepreneurship and creativity produced omnibus ecological validation and experience data. The study developed real-world qualitative and quantitative findings showing that exploring and curating with MI^{2}C, in comparison to using Google to explore and Word to curate, improves creativity and invention [Kerne et al. 2006, 2008; Kerne and Koh 2007].

9.5. Relevance on Par

Relevance has registered on par in both experiments that we have presented, serving as a barometer of validation for other metrics. We did not expect Relevance to increase, because we did not provide better search. If Relevance declined, it would indicate that increased Fluency, Flexibility, and Novelty were supported at the expense of maintaining topic focus. This was not the case. Thus, findings of Relevance on par add to the validity of the findings of the other ideation metrics.

9.6. Negative Correlation: Fluency and Visual Presentation

The decrease in Visual Presentation in the provocative stimuli condition, when mixed initiatives are used instead of direct-manipulation-only, is troubling. We interpret this result as a shortcoming not of MI^{2}C, per se, but of (1) an inherent tension in creative process and (2) issues with the interaction mechanisms of the present implementation.

9.6.1. Inherent Tension Between Fluency and Visual Presentation. Increased Fluency means more elements. More elements inherently result in higher cognitive and neuromuscular load. There is more to think about and more to manipulate, which is inherently more difficult. More elements makes Visual Presentation harder.

Further analysis of experimental data supports the hypothesis that Visual Presentation decreased as Fluency increased (Figure 11). We categorized compositions into four groups based on the number of clipping found objects curated in each. The dependent variable was mean Visual Presentation (0–3). The mean Visual Presentation was 1.69 for 0 to 20 elements (SE = 0.17), 1.365 for 21 to 40 elements (SE = 0.25), 1.04 for 41 to 60 elements (SE = 0.161), and 0.716 for 61 to 112 elements (SE = 0.212). One-way ANOVA revealed this variance in means to be statistically significant ($df = 1$, $F = 9.327$, $p < 0.004$).

At the same time, thinking about and manipulating more elements can play an essential role in the emergence of new ideas. Indeed, we have shown, even with the present implementation, concomitant increases in Emergence.

9.6.2. Visualization and Interaction Techniques Have the Potential to Ameliorate the Challenges of High Fluency. The present information composition apparatus only supports mouse-based interaction of individual elements in a finite 2D area. Based on this data, we hypothesize that the Visual Presentation of curators would be improved by giving them better ways to organize their compositions as the number of elements scales. Grouping mechanisms, like piles and layers, are hypothesized to significantly contribute to helping users develop coherent Visual Presentation, ameliorating the effects of increase in the number of elements. Further, techniques from information visualization, such as pannable, zoomable, and focus+context interfaces, are also hypothesized to benefit Visual Presentation of large collections.

[5] Instead of measuring the Web pages from which participants collected, this experiment counted the Variety of pages that they browsed. This measure of potential flexibility is a related process metric.

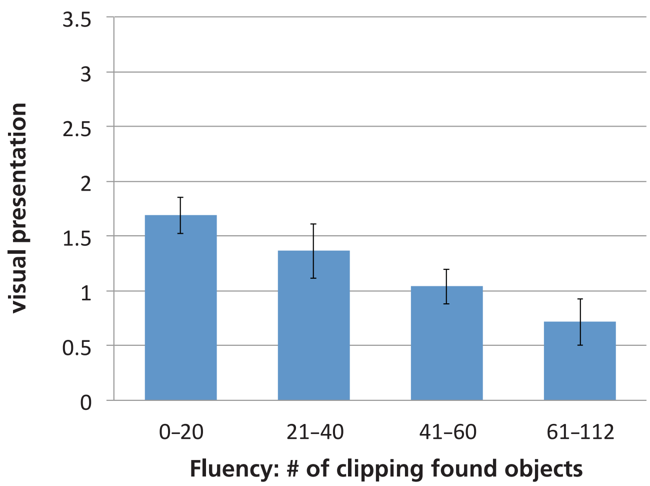

Fig. 11. As Fluency increases in curation products, mean Visual Presentation decreases. Increased Fluency means more elements. More elements are inherently more difficult to manipulate, organize, and design. Visual Presentation requires effective use of white space, which in a fixed size canvas inherently grows more scarce as Fluency increases.

Beyond the presentation medium, the mouse becomes an interaction bandwidth bottleneck when working with many elements. We hypothesize that richer, more embodied and higher bandwidth interaction, using multitouch and/or a stylus, will also serve, with the right interaction techniques and design, to mitigate the issues that make Visual Presentation difficult as the number of elements in a composition increases. Thus, future research can investigate hypotheses that integrating grouping mechanisms, nested spaces, and embodied interaction modalities will improve users' performance on IBI tasks.

10. IMPLICATIONS FOR DESIGN

This research develops information-based ideation activities as a locus of investigation and ideation metrics of curation as a means of evaluation. We develop implications for the design of evaluation methodology for investigating human activities involving information, which essentially involve developing ideas, and for IBI support environments and concomitant media of curation.

10.1. Information-Based Ideation Evaluation

We develop implications for the design of information-based ideation evaluation. The foundation is to engage people in open-ended IBI activities. Use study conditions to vary the tools, such as search or information visualization interfaces. Have participants develop curation products through this engagement in IBI. Then, in concert with mixed methods,[6] use ideation metrics like those presented in this article to evaluate the curations for each study condition. For metrics that require assessment by raters, specify clear criteria and procedures to promote consistency. For the Flexibility elemental ideation metric, we advocate computation over multiple dimensions.

10.1.1. Engage Users in IBI Tasks and Use Ideation Metrics to Assess Curation Products. Defining evaluations for sensemaking (e.g., Paul and Morris [2009], visual analytics [Thomas

[6]Complementary approaches include qualitative methods [Corbin and Strauss 2008], self-efficacy surveys [Carroll et al. 2009], and social and organizational studies [Amabile 1996].

and Cook 2006; Chang et al. 2009], and IBI support environments is challenging. Creative cognition researchers typically use quite artificial tasks to study insight problems (e.g., Mednick [1962]; Bowden et al. [2005]). IBI tasks are more connected to the real world. Whereas laboratory tasks and conditions do not correspond directly to real-world experiences, IBI laboratory tasks match more closely than prior simplistic sensemaking tasks (e.g., Woodruff et al. [2001]). For evaluating tools for supporting sensemaking and visual analytics, an IBI task can profitably be specified, in conjunction with an existing or new curation medium. Then, the use of ideation metrics of curation to evaluate creative products is prescribed.

Considering prior research, Paul and Morris [2009] employed what we can see is an IBI task to evaluate Co-Sense: “to explore the local metropolitan area and to come up with a joint plan for the weekend”. Co-Sense incorporates a curation tool for users to create answers to this task. Ideation metrics can be applied to resulting curations in order to compare Co-Sense with ad hoc methods that people typically use to search together.

Thiry et al. [2013] observed that users need to work with implicit structure in using timelines as a basis for curation of personally meaningful experiences. They articulated a goal to investigate how well the timeline framework supports creativity. The present IBI methodology can provide a basis for such research. “Stuff I’ve Seen” [Dumais et al. 2003] is another example of a system that could be evaluated in an interesting way with IBI tasks and metrics of curation on creative products.

10.1.2. Use IBI Tasks and Curation to Evaluate Information Visualization Systems. Diverse information visualization methods can potentially be used to generate provocative stimuli. Information visualization research has been seeking new evaluation methods (e.g., Bertini et al. [2008], Chang et al. [2009]). As Chang et al. [2009] observe, seminal researchers Card et al. [1999] declared that “the purpose of visualization is insight,” while Thomas and Cook [2005] likewise identified the purpose of visual analytics as to enable and discover insight. However, these researchers have surprisingly failed to align their definitions of insight with cognitive science [Chang et al. 2009].

The purpose of visualization is not insight, but ideation. Ideation refers to generating and exploring ideas, essential to brainstorming, divergent thinking, and insight problem solving [Finke et al. 1992]. The experimental study of ideation is characterized by open-ended tasks, but studies of insight have used closed-form problems, whose solutions are known in advance (e.g., Mednick [1962]; Bowden et al. [2005]). Ideation tasks can be designed in conjunction with information visualization systems and datasets to engage study participants toward new discoveries.

Comparatively evaluate an infovis system, for support in exploration of a dataset in comparison to a control condition, by having study participants engage in an IBI task, including curation of intermediate results. The curation medium and tool can be held constant while the exploration tool is varied. Apply ideation metrics to curation products to determine if one condition better supports creativity. The authors are making an extensible information composition environment freely available [Interface Ecology Lab 2013] to support such scientific research, as well as for end users in education and beyond.

10.1.3. Disambiguate Rating Rules. The derivation of holistic IBI metrics is presently based on the coordinated activities of expert raters. Variability in the reliability of measurement is an inherent problem for methods involving psychological metrics [Nunnally and Bernstein 1994]. Disambiguating the rules for rating is an essential method for improving measurement [Nunnally and Bernstein 1994]. Clear definition of what is being measured is necessary for investigation of significant phenomena.

Ambiguity across raters produces more random ratings and thus reduces the accuracy of measuring significant effects of an IBI support environment.

To disambiguate ideation metrics' ratings, metrics should first be defined to be distinct—as mutually independent as possible. Otherwise, raters will become confused and sometimes award the same point in multiple places for a single accomplishment by a study participant who curated an experimental product. Likewise, clearly specify rating levels for each metric to reduce ambiguity and confusion. Raters should meet to calibrate their interpretations of ratings specifications in context. The round-down rule should be consistently applied: assign the lower score for partial or ambiguous attainment of a metric rating. These principles serve as the basis for high levels of inter-rater reliability in experimental procedures.

We achieved inter-rater reliability correlation coefficients of 0.85 and 0.83 using the holistic ideation metrics of the curation rating procedure. By comparison, on an emotion-labeling task, Snow et al. [2008] achieved Pearson's correlation coefficients of 0.58 among expert raters and .43 when comparing nonexperts to experts. A foundational text on psychological measurement prescribes that for construct validation research, reliability of .70 is quite adequate, and that "increasing reliabilities beyond .80 in basic research is wasteful" [Nunnally and Bernstein 1994]. Future research can investigate invoking holistic ideation metrics and methods to align raters using crowdsourcing, with the goal of scaling alignment of ratings to larger datasets. Prior methods, such as providing external feedback to microtask workers [Dow et al. 2012], are likely to be involved.

10.1.4. Measure Flexibility in Multiple Dimensions. The Flexibility ideation metric of curation addresses the diversity of approaches and positions that a participant engages in performing an IBI task. Ideational Flexibility has been constituted as the number of categories of ideas that participants assemble [Shah et al. 2002]. We need to discover diverse mechanisms for investigating and measuring Flexibility in IBI. Categories canbe construed in many ways. We presently measure Flexibility based on from where users collect information, such as the source documents, Web sites, and types of Web sites. This is a beginning. Flexibility can also be addressed conceptually. We propose that future research additionally measure distinct forms of conceptual Flexibility by labeling answers with concepts, using ontologies (e.g., Rector et al. [1997]), facets (e.g., Hearst [2008]), social media, and visual features (e.g., Snoek and Worring [2009]). The uniqueness of such labelings, in the context of an experiment, will also form a basis for new Novelty metrics.

10.2. Curation Media and IBI Support Environments

We develop implications for the design of curation media and concomitant IBI support environments, proceeding in granularity from the representation of whole collections to smaller units. Support environments constitute affordances and media for curation. Engagement in IBI involves searching, choosing digital found objects, annotating, reflecting, and synthesizing as a basis for creating new ideas. More generally, we identify a need to connect disciplines in the investigation of creativity support environments.

10.2.1. Support Curation to Bridge the Synthesis Gap. Whether users are exploring a space of ideas through information visualization or through Web browsing, when their goal is to develop new ideas, they need to collect and connect intermediate results. In this direction, Blake and Pratt [2006] developed a collaborative information synthesis model. It includes "hypothesis projection" without more general consideration of ideation and creative cognition. Beyond their prescription, much of the model is equally as relevant for individuals as it is for research teams.

As per the prescription of visual analytics researchers, there is an urgent need to develop capabilities for synthesizing information across sources. Choosing, assembling, annotating, and reflecting on—that is, *synthesizing* across digital objects—is essential to curation. Curation with the goal of developing new ideas is the crux of IBI. As we have shown (Section 5.3), epistemologically, *synthesis* goes beyond "analysis." Thus, to address user needs, we need to build research beyond the conceptual confines of visual analytics by supporting engagement in IBI in conjunction with analytics (e.g., in decision-making contexts). This will result in better human-centered computing.

10.2.2. Integrate Implicit and Explicit Structure. In the present research, explicit structure is represented in the interface as metadata while internally, in the agents' model, as the semantic associationality of clippings within documents and connecting hyperlinks. In other information visualization systems, formal structure means the details of how an infovis view is constituted, including what objects have focus, what are in context, and, in the spirit of the mantra of Shneiderman [1996], the level of zoom, the settings of filters, and relevant details-on-demand.

The findings of Marshall and Rogers [1992] on the importance of support for implicit structure are two decades old. Researchers continue to rediscover the tension between explicit and implicit structure. Thiry et al. [2013] found that although a timeline supported people in organizing life experiences, the requirement that events must be chronologically organized became a limitation, as users remember the dates of some significant events but forget others. Our findings about the value of curation to bridge the synthesis gap support building tools that integrate implicit and explicit structure. Formal structure helps organize information. It can be used to support people discovering associations through models. Yet, by enabling free association, informal structure helps people synthesize information and form emergent ideas. The juxtaposition of less expected, more remote thoughts is a significant aspect of creativity [Wilkenfeld and Ward 2001].

10.2.3. Tightly Couple Curation with Visualization. We have shown that integrating information visualization with curation significantly increases ideation, and that users overwhelmingly prefer it across aspects of their experiences on IBI tasks. When used as provocative stimuli, MI^{2}C's integration also promotes Emergence, the ideation metric associated with synthesis. Thus, coupling curation with visualization addresses urgent needs articulated by Thomas and Cook's visual analytics agenda, with scope *beyond analytics*, to bridge the synthesis gap.

Information visualization systems will benefit from directly addressing how people use the insights they gain through interaction with a visualization. Interacting with a visualization is an inherently high cognitive load activity. Users need fluid means for collecting and reflecting on insights they gain en route to reduce the cognitive and neuromuscular load of switching contexts to copy, paste, and annotate intermediate results. They need support in synthesis of intermediate results. Spatial integration is one aspect of this bridging of the synthesis gap. Another, semantic coupling, would enable users to seamlessly bring visual clippings with semantic metadata over from a visualization to a curation space, to help them maintain context, and later index back to intermediate results in the visualization environment.

10.2.4. Use Visual Clipping Found Objects as a Medium for Elements of Curation. Visual clipping found objects directly and simply represent ideas. They afford refinding. They resemble predigital 3×5 cards, but are more flexible and powerful. They range in size from a phrase to sentences to an image.

Clipping images overcomes the text-centric neglect of visual cognition in working memory [Baddeley 1986], which characterizes prior approaches to the medium of elements of a large collection (e.g., Pirolli et al. [1996]). A study participant reported that information composition's small elements constituted 'everything you wanted to know" [Koh et al. 2007]. The findings of Teevan et al. [2009] with regard to visual snippets are consistent. However, they derived only a single, fixed visual snippet per Web page. We rather prescribe, for documents with more than one meaningful image, enabling users to choose clippings based on task context. The success of Pinterest, which gives the user this choice when she curates a Web page, supports this implication. Curating heterogeneous clippings enables representing diverse points of view. IBI support environment builders should design tools for collecting clipping found objects that enable users to capture small chunks of information essential to their activities and experiences.

10.2.5. Contextualize Clipping Found Objects with Metadata Semantics. Clippings, by nature, are contextualized. They are contextualized by the documents in which they are found. They are contextualized by the activities during which they are curated. In IBI, the context can prove as significant as the clipped material itself. As an element of curation, a little clipping can possess a lot of context.

As a medium for elements of curation, clippings function as referential digital found objects conjoined with metadata that enables refunding source documents and visualization contexts. When the clipped material is an image, this results in a visual bookmark, whether the source is a Web page or a Tableau analytics view.

Metadata semantics can go levels further in contextualizing clippings as digital found objects. Metadata can show users the article in a digital library or news feed that a clipping represents. They can show what product a clipping is, how much it costs, how it has been reviewed, and how to buy it. Visual semantic clipping found objects can show a document's references to prior articles or patents or films, as well as articles that cite it or actors who were in it. They can incorporate all parameters that characterize an information visualization view, affording return to deep analytic contexts. As elements of curation in social media, visual semantic clippings can show not just how people have tagged it, but also what *your* friends and colleagues said about the object at hand. In all manner of curation tools, including but not limited to social networks such as Pinterest and Facebook, and for information visualization systems, contextualize visual semantic clipping found objects, as elements of curation, in social and citation graphs of referentiality.

10.2.6. Creative Cognition Basis for Creativity Support Environments. Using cognitive psychologyas a basis for modeling interaction and experience is as old as the field of HCI itself. Card et al. [1980] drew from psychological research on "routine cognitive skill." The present research draws on almost a century of experimental creative cognition research (e.g., Maier [1931]; Guilford [1950]; Finke et al. [1992]; Sternberg [1999]).

We advocate creative cognition as a basis for designing interactive creativity support environments and their evaluation:

(1) Creative ideas result from cognitive processes.[7]
(2) Design interactive system features to support creativity based on the aspects of cognition that they are intended to support.
(3) Use experimental methods, such as the IBI evaluation methodology, to evaluate individual and integrated interactive system components.
(4) Use open-ended divergent thinking tasks, such as IBI tasks, to study tools designed to support creative cognition.

[7]They also result from emotional and social processes.

10.2.7. Use Art to Conceptualize, Science to Validate. With the ubiquitous proliferation of computing, beyond the prognostication of Weiser and Seely Brown [1997], HCI is enmeshed in society. As such, HCI is as much a form of culture as it is a science.

The province of innovation in culture is art. From Gutenberg to Rauschenberg to Lady Gaga and Jay Z, the hallmarks of media are the expressive forms that people create with them. Technology's role is to build enabling platforms. Thus, look to art for conceptual basis. The paradigm of mixed-initiative information composition is based on a notion of curation of found objects introduced by Marcel Duchamp, and one of composition of found objects advanced by John Cage.

In this vein, Benford et al. [2011] observed that "arts-based research can explore ways in which new technologies . . . change how people interact with one another."

Whereas validation in art is quirky, subject to personalities and trends, systematic evaluation is a hallmark of science. Thus, creativity support environments tend toward art-science hybrids. Although such hybridization is neither necessary nor sufficient per se, it intuitively makes sense for the success of creativity support environments.

The empirical aspect of the present research does not conflict with a sense of the importance of the role of intuition in creativity and its support. We thus add grist to the STEM to STEAM [2013] movement for including art in a pantheon of fields, along with science, technology, engineering, and math, because artistic methodologies hold keys to innovation.

11. CONCLUSION

In this article, we developed the paradigm of IBI and a methodology for comparative evaluation of IBI support environments consisting of the following: (1) engage people performing IBI tasks, (2) elicit their curation products, (3) measure creativity with ideation metrics of curation, and (4) compare ideation metrics distributions across conditions.

Prior notions, such as sensemaking, information foraging, and exploratory search, are part of IBI. Information visualization tools can likewise be involved in IBI experiences. Why bother with this new terminology given the prior research paradigms?

IBI shifts framing to a human-centered perspective focusing on creative processes that are meaningful to people as they engage with information. IBI provides methods for evaluation in the form of metrics of curation that can be applied to differentially measure the impact of support environments on aspects of creativity. *Creative cognition provides means for investigation of IBI support based on research on mechanisms of the mind.* At the same time, creative cognition does not make the epistemological error of claiming to account for all aspects of creativity [Finke et al. 1992]. Neither does the present research. Other approaches from the arts and sciences are necessary. For example, qualitative and social investigations of IBI activities are needed.

We posited curation as the form of products that people create through engagement in IBI tasks. That is, collecting digital objects, assembling them, annotating, and reflecting. We note here that curation is in fact an important part of any report writing process. "Annotation" can involve any amount of writing, sketching, and recording. The terminology emphasizes the associationality afforded by the Internet for connecting digital found objects. Digital curation denotes a rich form of report and presentation with abundant media and citations. Digital objects referenced in curation encompass those that are self-made and those that are readymade. These elements of curation may be stored in personal, family, organizational, institutional, and national repositories. The mini-c digital objects that people make just for themselves and the big data of large-scale publications, datasets, data warehouses, and statistics are all at play.

We present information composition as an expressive and creative medium of curation for people engaged in IBI tasks. Composition emphasizes juxtaposing and blending found objects and annotation. Composition affords synthesis and idea generation, supporting people in thinking across intermediate findings of in-depth analytics.

We developed mixed-initiative information composition, in which software agents engage in information retrieval and visualization over time, in collaboration with the user, to afford exploration and curation in the performance of IBI tasks. We conducted laboratory studies using the IBI methodology to evaluate MI^2C. We found that MI^2C functions as a source of provocative stimuli that help people overcome fixation and become more creative on IBI tasks. The integration of information retrieval and visualization agents resulted in increased Fluency, Flexibility, Novelty, and Emergence. We also found that MI^2C bridges the synthesis gap. Spatial and semantic coupling of curation and visualization tools resulted in increased Fluency, Flexibility, and Novelty.

The creation of new media for curation and concomitant interactive environments that support IBI remains an open field. The present case study of information composition constitutes a signpost on an emerging terrain. Building interactive and mixed-initiative information environments that effectively provoke and support creativity is a long-term research agenda, which requires sustained, integrated investigation of art, design, cognition, sociology, information visualization, information semantics and retrieval, sensing and recognition, interaction techniques, and digital media, including componentized evaluation methods that investigate the efficacy of system components. The IBI methodology gives researchers empirical, objective means for evaluating new curation media and IBI support environments.

Others have evaluated creativity support through subjective experience measures (e.g., Yee et al. [2003]), factor-based self-assessment of creativity experiences [Carroll et al. 2009], social psychology methods [Tripathi and Burleson 2012], qualitative analysis (e.g., Hornecker et al. [2008]), case studies (e.g., Perer and Shneiderman [2008]), and advertisement click-through [Dow et al. 2010]. *By measuring the levels of components of creativity in participants' curations, amidst the panoply of methods, IBI can play an important role in our understanding of which tools actually support creative engagement with information, and how.* This extension of creative cognition into HCI contributes deeper understanding of specific mechanisms of creativity and rigorous evaluation of how interactive systems support them. Information-based ideation has the potential to help researchers unleash the power of data, big and small, to fuel invention that drives our economies and infuses our personal lives with satisfaction.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation (NSF) under grants IIS-074742 and IIS-1247126. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. Further support was provided by a sabbatical grant to the first author by the University of Nottingham's Horizon Institute of Digital Economy.

We thank our reviewers and editor for their time and effort to work with us to strengthen this article. Thanks to Reviewer 1 for suggesting developing a Flexibility metric based on source Web site type. Thanks to Charlotte Lee and Stuart Reeves for thoughtful feedback on drafts. Thanks to statistics luminary Emanuel Parzen for advice on methods for computing inter-rater reliability. Thanks to the first author's wife, Shannon Taylor-Kerne, for sustained patience and support.

REFERENCES

Teresa M. Amabile. 1996. Creativity and innovation in organizations. *Harvard Business School Background Note* 396–239.

Robert Amar and John Stasko. 2004. BEST PAPER: A knowledge task-based framework for design and evaluation of information visualizations. In *Proceedings of the IEEE Symposium on Information Visualization*. IEEE Computer Society, Washington, DC, 143–150. DOI:http://dx.doi.org/10.1109/INFOVIS.2004.10

Doug Babb. 2012. Shrinking World of World News. *Managing Issues: CFM Public Affairs Blog.* Available at http://www.cfm-online.com/public-affairs-blog/2012/6/25/shrinking-world-of-world-news.html.

Alan Baddeley. 1986. *Working Memory*. Clarendon Press.

Marcia J. Bates. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review* 13, 5, 407–424.

Ronald A. Beghetto and James C. Kaufman. 2007. Toward a broader conception of creativity: A case for "mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts* 1, 2, 73–79.

Nicholas J. Belkin, Robert N. Oddy, and Helen M. Brooks. 1982. ASK for information retrieval. Part I: Background and theory. *Journal of Documentation* 38, 2, 61–71.

Steve Benford, Andy Crabtree, Martin Flintham, Chris Greenhalgh, Boriana Koleva, Matt Adams, Nick Tandavanitj, Ju Row Farr, Gabriella Giannachi, and Irma Lindt. 2011. Creating the spectacle: Designing interactional trajectories through spectator interfaces. *ACM Transactions on Computer-Human Interacttion* 18, 3, Article 11. DOI:http://dx.doi.org/10.1145/1993060.1993061

Enrico Bertini, Adam Perer, Catherine Plaisant, and Guiseppe Santucci. 2008. Beyond time and errors: Novel evaluation methods for information visualization. In *Proceedings of CHI'08: Extended Abstracts on Human Factors in Computing Systems CHI EA'08).* 3913–3916.

Catherine Blake and Wanda Pratt. 2006. Collaborative information synthesis II: Recommendations for information systems to support synthesis activities. *JASIST* 57, 14, 1888–1895.

Margaret Boden. 2003. *The Creative Mind: Myths and Mechanisms*. Routledge.

Edward M. Bowden, Mark Jung-Beeman, Jessica Fleck, and John Kounios. 2005. New approaches to demystifying insight. *Trends in Cognitive Sciences* 9, 7, 322–328.

Bill Buxton. 2007. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann, San Francisco, CA.

John Cage. 1961. *Silence: Lectures and Writings*. Wesleyan University Press.

Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. 1999. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann. Available at http://books.google.com/books?id=wdh2gqWfQmgC.

Stuart K. Card, Thomas P. Moran, and Allen Newell. 1980. Computer text-editing: An information-processing analysis of a routine cognitive skill. *Cognitive Psychology* 12, 1 (January 1980), 32–74.

Erin A. Carroll, Celine Latulipe, Richard Fung, and Michael Terry. 2009. Creativity factor evaluation: Towards a standardized survey metric for creativity support. In *Proceedings of the 7th ACM Conference on Creativity and Cognition (C&C'09)*. ACM, New York, NY, 127–136. DOI:http://dx.doi.org/10.1145/1640233.1640255

Remco Chang, Caroline Ziemkiewicz, Tera Marie Green, and William Ribarsky. 2009. Defining insight for visual analytics. *IEEE Computer Graphics and Applications* 29, 2 (March 2009), 14–17. DOI:http://dx.doi.org/10.1109/MCG.2009.22

Carina Chocano. 2012. Pinterest, tumblr and the trouble with 'curation.' *New York Times Magazine.*

Bo T. Christensen and Christian D Schunn. 2007. The relationship between analogical distance to analogical function and pre-inventive structure: The case of engineering design. *Memory Cognition* 35, 29–38.

Juliet Corbin and Anselm Strauss. 2008. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage.

Council on Competitiveness. 2005. *Innovate America: National Innovation Initiative Summit and Report*.

Steven Dow, Anand Kulkarni, Scott Klemmer, and Björn Hartmann. 2012. Shepherding the crowd yields better work. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW'12)*. ACM, New York, NY, 1013–1022. DOI:http://dx.doi.org/10.1145/2145204.2145355

Steven P. Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L. Schwartz, and Scott R. Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interactions* 17, 4, Article 18. DOI:http://doi.acm.org/10.1145/1879831.1879836

Susan Dumais, Edward Cutrell, JJ Cadiz, Gavin Jancke, Raman Sarin, and Daniel C. Robbins. 2003. Stuff I've seen: A system for personal information retrieval and re-use. In *Proceedings of the 26th Annual International ACM SIGR Conference on Research and Development in Information Retrieval*. ACM, New York, NY, 72–79. DOI:http://dx.doi.org/10.1145/860435.860451

Clive Dym, Alice Agogino, Ozgur, Daniel Frey, and Larry Leifer. 2005. Engineering design thinking, teaching, and learning. *Journal of Engineering Communication* 94, 103–120.

Ronald A. Finke, Thomas B. Ward, and Steven M. Smith. 1992. *Creative Cognition: Theory, Research, and Applications*. MIT Press. Available at http://books.google.com/books?id=FVOwHAAACAAJ.

William W. Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'03)*. ACM, New York, NY, 233–240. DOI:http://dx.doi.org/10.1145/642611.642653

Arthur M. Glenberg and William E. Langston. 1992. Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language* 31, 2, 129–151. DOI:http://dx.doi.org/DOI: 10.1016/0749-596X(92)90008-L

Joy P. Guilford. 1950. Creativity. *American Psychologist* 5, 444–454.

Joy P. Guilford. 1956. The structure of intellect. *Psychological Bulletin* 53, 267–293.

Joy P. Guilford. 1968. *Intelligence, Creativity, and Their Educational Implications*. Knapp.

Joshua Hailpern, Erik Hinterbichler, Caryn Leppert, Damon Cook, and Brian P. Bailey. 2007. TEAM STORM: Demonstrating an interaction model for working with multiple ideas during creative group work. In *Proceedings of the ACM SIGCHI Conference on Creativity and Cognition*. 193–202. DOI:http://dx.doi.org/10.1145/1254960.1254987

James A. Hampton. 1997. Emergency attributes in combine concepts. In *Creative Thought: An Investigation of Conceptual Structures and Processes*. American Psychological Association, 83–110.

Harvard College Writing Program. 2011. *Writing Resources*. *The Writing Center*. Available at http://isites.harvard.edu/icb/icb.do?keyword=k33202&pageid=icb.page143936.

William L. Hays. 1988. *Statistics* (4th ed.). Holt, Rinehart and Winston.

Marti A. Hearst. 2008. UIs for faceted navigation: Recent advances and remaining open problems. In *Proceedings of the Workshop on Computer Interaction and Information Retrieval (HCIR'08)*.

Yosef Hochberg and Ajit C. Tamhane. 1987. *Multiple Comparison Procedures*. Wiley.

Eva Hornecker, Paul Marshall, Nick Sheep Dalton, and Yvonne Rogers. 2008. Collaboration and interference: Awareness with mice or touch input. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW'08)*. ACM, New York, NY, 167–176. DOI:http://dx.doi.org/10.1145/1460563.1460589

Eric Horvitz. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: The CHI Is the Limit (CHI'99)*. ACM, New York, NY, 159–166. DOI:http://dx.doi.org/10.1145/302979.303030

Ivan Illich. 2001. *Tools for Conviviality*. Marion Boyars.

Interface Ecology Lab. 2013. IdeaMâché. Retrieved May 22, 2014, from http://ideamache.ecologylab.net.

David G. Jansson and Steven M. Smith. 1991. Design fixation. *Design Studies* 12, 1, 3–11. DOI:http://dx.doi.org/DOI: 10.1016/0142-694X(91)90003-F

Benny Johnson. 2013. 24 Delightful Inauguration Firsts. *BuzzFeed*. Retrieved May 22, 2013, from http://www.buzzfeed.com/bennyjohnson/23-delightful-inauguration-firsts.

William Jones, Susan Dumais, and Harry Bruce. 2002. Once found what then? A study of "keeping" behaviors in the personal use of Web information. In *Proceedings of the 65th ASIST Annual Meeting (ASIST'02)*. 391–402.

Wassily Kandinsky. 1994. *Kandinsky, Complete Writings on Art*. Da Capo Press.

James C. Kaufman and Robert J. Sternberg (Eds.). 2010. *The Cambridge Handbook of Creativity*. Cambridge University Press, Cambridge, MA.

Andruid Kerne and Eunyee Koh. 2007. Representing collections as compositions to support distributed creative cognition and situated creative learning. *New Review of Hypermedia and Multimedia* 13, 2 (December 2007), 135–162. DOI:http://dx.doi.org/10.1080/13614560701711859

Andruid Kerne, Eunyee Koh, Blake Dworaczyk, J. Michael Mistrot, Hyun Choi, Steven M. Smith, Ross Graeber, Daniel Caruso, Andrew Webb, Rodney Hill, and Joel Albea. 2006. combinFormation: A mixed-initiative system for representing collections as compositions of image and text surrogates. In *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'06)*. ACM, New York, NY, 11–20. DOI:http://dx.doi.org/10.1145/1141753.1141756

Andruid Kerne, Eunyee Koh, Steven Smith, Hyun Choi, Ross Graeber, and Andrew Webb. 2007. Promoting emergence in information discovery by representing collections with composition. In *Proceedings of the 6th ACM SIGCHI Conference on Creativity and Cognition (C&C'07)*. ACM, New York, NY, 117–126. DOI:http://dx.doi.org/10.1145/1254960.1254977

Andruid Kerne, Eunyee Koh, Steven M. Smith, Andrew Webb, and Blake Dworaczyk. 2008. combinFormation: Mixed-initiative composition of image and text surrogates promotes information discovery. *ACM Transactions on Information Systems* 27, 1 (December 2008), Article 5. DOI:http://dx.doi.org/10.1145/1416950.1416955

Andruid Kerne, Yin Qu, Andrew M. Webb, Sashikanth Damaraju, Nic Lupfer, and Abhinav Mathur. 2010. Meta-metadata: A metadata semantics language for collection representation applications.

In *Proceedings of the 19th ACM Iinternational Conference on Information and Knowledge Management (CIKM'10)*. ACM, New York, NY, 1129–1138. DOI:http://dx.doi.org/10.1145/1871437.1871580

Andruid Kerne, Steven M. Smith, Eunyee Koh, Hyun Choi, and Ross Graeber. 2008. An experimental method for measuring the emergence of new ideas in information discovery. *International Journal of Human-Computer Interaction* 24, 460–477. DOI:http://dx.doi.org/10.1080/10447310802142243

Mi Jeong Kim and Mary Lou Maher. 2005. Comparison of designers using a tangible user interface and a graphical user interface and the impact on spatial cognition. In *Proceedings of Human Behaviour in Design '05*, J. S. G. U. Lindemann and U. Lindemann (Eds.). University of Sydney.

Eunyee Koh, Andruid Kerne, Andrew Webb, Sashikanth Damaraju, and David Sturdivant. 2007. Generating views of the buzz: Browsing popular media and authoring using mixed-initiative composition. In *Proceedings of the 15th International Conference on Multimedia (MULTIMEDIA'07)*. ACM, New York, NY, 228–237. DOI:http://dx.doi.org/10.1145/1291233.1291282

Nicholas W. Kohn and Steven M. Smith. 2009. Partly versus completely out of your mind: Effects of incubation and distraction on resolving fixation. *Journal of Creative Behavior* 43, 2, 102–118.

Nicholas W. Kohn and Steven M. Smith. 2011. Collaborative fixation: Effects of others' ideas on brainstorming. *Applied Cognitive Psychology* 25, 3, 359–371. DOI:http://dx.doi.org/10.1002/acp.1699

Rosalind Krauss. 1998. *The Optical Unconscious*. MIT Press.

Ranjitha Kumar, Jerry O. Talton, Salman Ahmad, and Scott R. Klemmer. 2011. Bricolage: Example-based retargeting for Web design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*. ACM, New York, NY, 2197–2206. DOI:http://dx.doi.org/10.1145/1978942.1979262

Rhema Linder, Clair Snodgrass, and Andruid Kerne. 2014. Everyday ideation: All of my ideas are on Pinterest. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*. ACM, New York, NY, 2411–2420.

Julie S. Linsey, Matthew G. Green, J. T. Murphy, Kristin L. Wood, and Art B. Markman. 2005. Collaborating to success: An experimental study of group idea generation techniques. In *Proceedings of the ASME Design Theory and Methodology Conference.* 24–28.

Lucy Lippard. 1972. *Dadas on Art*. Prentice Hall.

Kate Maccariello. 2013. The World Is My Oyster. *Pinterest.* Retrieved May 22, 2014, from http://www.pinterest.com/katemaccariello/havelovewilltravel/.

Norman R. F. Maier. 1931. Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology* 12, 2, 181–194. DOI:http://dx.doi.org/DOI:10.1037/h0071361

Gary Marchionini. 2006. Exploratory search: From finding to understanding. *Communications of the ACM* 49, 4 (April 2006), 41–46. DOI:http://dx.doi.org/10.1145/1121949.1121979

Joe Marks, Brad Andalman, Paul A. Beardsley, William Freeman, Sarah Gibson, Jessica, Thomas Kang, Brian Mirtich, Hanspeter Pfister, Wheeler Ruml, Kathy Ryall, Joshua, and Stuart Shieber. 1997. Design Galleries: A general approach to setting parameters for computer graphics and animation. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'97)*. ACM Press/Addison-Wesley, New York, NY, 389–400. DOI:http://dx.doi.org/10.1145/258734.258887

Catherine C. Marshall and Sara Bly. 2005. Saving and using encountered information: Implications for electronic periodicals. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'05)*. ACM, New York, NY, 111–120. DOI:http://dx.doi.org/10.1145/1054972.1054989

Catherine C. Marshall and Russell A. Rogers. 1992. Two years before the mist: Experiences with Aquanet. In *Proceedings of the ACM Conference on Hypertext (ECHT'92)*. ACM, New York, NY, 53–62. DOI:http://dx.doi.org/10.1145/168466.168490

Catherine C. Marshall and Frank M. Shipman III. 1993. Searching for the missing link: Discovering implicit structure in spatial hypertext. In *Proceedings of the 5th ACM Conference on Hypertext (HYPERTEXT '93)*. ACM, New York, NY, 217–230. DOI:http://dx.doi.org/10.1145/168750.168826

Sarnoff A. Mednick. 1962. The associative basis of the creative process. *Psychological Review*, 220–232.

Pranav Mistry. 2012. SixthSense—a wearable gestural interface. Video. Retrieved May 22, 2014, from http://www.pranavmistry.com/projects/sixthsense/.

National Academy of Engineering. 2005. *Engineering Research and America's Future: Meeting the Challenges of a Global Economy*. National Academies Press.

National Academy of Engineering. 2010. *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5*. National Academies Press.

Brent A. Nelson, Jamal O. Wilson, David Rosen, and Jeannette Yen. 2009. Refined metrics for measuring ideation effectiveness. *Design Studies* 30, 6, 737–743.

Chris North and Ben Shneiderman. 2000. Snap-together visualization: Can users construct and operate coordinated visualizations? *International Journal of Human-Computer Studies* 53, 5, 715–739. DOI:http://dx.doi.org/10.1006/ijhc.2000.0418

Jum C. Nunnally and Ira H. Bernstein. 1994. *Psychometric Theory*. Issue 972. McGraw-Hill Series in Psychology. Available at http://books.google.com/books?id=r0fuAAAAMAAJ.

OpenDNS. 2013. OpenDNS Community—Domain Categories. Available at http://community.opendns.com/domaintagging/categories.

Alex F. Osborn. 1963. *Applied Imagination: Principles and Procedures of Creative Problem-Solving*. Scribner. Available at http://books.google.com/books?id=2GcaAAAAIAAJ.

Oxford University Press. 2013. *Oxford English Dictionary*. Available at http://www.oed.com.

Emmanuel Parzen. 1960. *Modern Probability Theory and Its Applications*. Wiley.

Emanuel Parzen and Subhadeep Mukhopadhyay. 2012. Modeling, dependence, classification, united statistical science, many cultures. *arXiv preprint arXiv:1204.4699*.

Sharoda A. Paul and Meredith Ringel Morris. 2009. CoSense: Enhancing sensemaking for collaborative Web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*. ACM, New York, NY, 1771–1780. DOI:http://dx.doi.org/10.1145/1518701.1518974

Adam Perer and Ben Shneiderman. 2008. Integrating statistics and visualization: Case studies of gaining clarity during exploratory data analysis. In *Proceedings of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems (CHI'08)*. ACM, New York, NY, 265–274. DOI:http://dx.doi.org/10.1145/1357054.1357101

Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological Review* 106, 4, 643–675.

Peter Pirolli, Patricia Schank, Marti Hearst, and Christine Diehl. 1996. Scatter/gather browsing communicates the topic structure of a very large text collection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'96)*. ACM, New York, NY, 213–220. DOI:http://dx.doi.org/10.1145/238386.238489

Douglas Quenqua. 2013. The Boy Wonder of BuzzFeed. *New York Times*. Retrieved May 22, 2014, from http://www.nytimes.com/2013/02/17/fashion/ben-smith-the-boy-wonder-of- buzzfeed.html.

Alan L. Rector, Sean Bechhofer, Carole A. Goble, Ian Horrocks, W. Anthony Nowlan, and W. Danny Solomon. 1997. The GRAIL concept modelling language for medical terminology. *Artificial Intelligence in Medicine* 9, 2, 139–171. DOI:http://dx.doi.org/DOI: 10.1016/S0933-3657(96)00369-7

Daniel Ritchie, Ankita Arvind Kejriwal, and Scott R. Klemmer. 2011. d.tour: Style-based exploration of design example galleries. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST'11)*. ACM, New York, NY, 165–174. DOI:http://dx.doi.org/10.1145/2047196.2047216

Steven Rosenbaum. 2011. *Curation Nation: How to Win in a World Where Consumers Are Creators*. McGraw-Hill.

Mark A. Runco and Robert S. Albert. 1985. The reliability and validity of ideational originality in the divergent thinking of academically gifted and nongifted children. *Educational and Psychological Measurement* 45, 3, 483–501.

Mark A. Runco and Shawn M. Okuda. 1988. Problem discovery, divergent thinking, and the creative process. *Journal of Youth and Adolescence* 17, 3, 211–220. DOI:http://dx.doi.org/10.1007/BF01538162 10.1007/BF01538162

Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. 1993. The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems (CHI'93)*. ACM, New York, NY, 269–276. DOI:http://dx.doi.org/10.1145/169059.169209

Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Information Processing and Management* 24, 5, 513–523. DOI:http://dx.doi.org/DOI: 10.1016/0306-4573(88)90021-0

Colleen M. Seifert, David E. Meyer, Natalie Davidson, Andrea L. Patalano, and Ilan Yaniv. 1995. Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In *The Nature of Insight*. MIT Press.

Jami J. Shah. 1998. Experimental investigation of collaborative techniques for progressive idea generation. In *Proceedings of ASME Design Theory*. 50.

Jami J. Shah, Steve M. Smith, and Noe Vargas-Hernandez. 2002. Metrics for measuring ideation effectiveness. *Design Studies* 24, 2, 111–134. DOI:http://dx.doi.org/DOI: 10.1016/S0142-694X(02)00034-0

Jami J. Shah, Steve M. Smith, Noe Vargas-Hernandez, David R. Gerkins, and Muqi Wulan. 2003. Empirical studies of design ideation: Alignment of design experiments with lab experiments. In *Proceedings of the ASME Conference on Design Theory and Methodology*. 1–10.

Jami J. Shah, Noe Vargas-Hernandez, Joshua D. Summers, and Santosh Kulkarni. 2001. Collaborative sketching (C-Sketch): An idea generation technique for engineering design. *Journal of Creative Behavior* 35, 3, 168–198.

Ben Shneiderman. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages (VL'96)*. IEEE Computer Society, Washington, DC, 336–. Available at http://dl.acm.org/citation.cfm?id=832277.834354.

Ben Shneiderman, Gerhard Fischer, Mary Czerwinski, Brad Myers, and Mitch Resnick. 2005. *Creativity Support Tools: A Workshop Sponsored by the National Science Foundation*.

Herbert A. Simon. 1960. *The New Science of Management Decision*. Harper & Brothers, 50–xii.

Dean Keith Simonton. 1999. Creativity as blind variation and selective retention: Is the creative process darwinian? *Psychological Inquiry* 10, 4, 309–328. Available at http://www.jstor.org/stable/1449455.

Tobias Skog, Sara Ljungblad, and Lars Erik Holmquist. 2003. Between aesthetics and utility: Designing ambient information visualizations. In *Proceedings of the IEEE Symposium on Information Visualization*. 233–240.

Smith M. Smith. 1996. Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In *The Nature of Insight*, R. J. Sternberg and J. E. Davidson (Eds.). MIT Press, 229–251.

Steven M. Smith and Edward Vela. 1991. Incubated reminiscence effects. *Memory and Cognition* 19, 2, 168–176. DOI:http://dx.doi.org/10.3758/BF03197114

Steven M. Smith, Thomas B. Ward, and Ronald A. Finke (Eds.). 1995. *The Creative Cognition Approach*. MIT Press.

Steven M. Smith, Thomas B. Ward, and Ronald A. Finke. 1995. *The Creative Cognition Approach*. MIT Press. Available at http://books.google.com/books?id=iEMWI4Z1rF8C.

Steven M. Smith, Thomas B. Ward, and Jay Schumacher. 1993. Constraining effects of examples in a creative generation task. *Memory and Cognition* 21, 6, 837–845. DOI:http://dx.doi.org/10.3758/BF03202751

Steven M. Smith and Steven E. Blankenship. 1989. Incubation effects. *Bulletin of the Psychonomic Society* 27, 4, 311–314.

Steven M. Smith and Steven E. Blankenship. 1991. Incubation and the persistence of fixation in problem solving. *American Journal of Psychology* 104, 1, 61–87. Available at http://www.jstor.org/stable/1422851.

Steven M. Smith and Deborah R. Tindell. 1997. Memory blocks in word fragment completion caused by involuntary retrieval of orthographically related primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23, 2, 355–370.

Cees G. M. Snoek and Marcel Worring. 2009. Concept-based video retrieval. *Foundations and Trends in Information Retrieval* 4, 2, 215–322.

Rion Snow, Brendan O'Connor, Daniel Jurafsky, and Andrew Y. Ng. 2008. Cheap and fast—but is it good?: Evaluating non-expert annotations for natural language tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP'08)*. Association for Computational Linguistics, Stroudsburg, PA, 254–263. Available at http://dl.acm.org/citation.cfm?id=1613715.1613751.

Shuang Song and Alice M. Agogino. 2004. Insights on designers' sketching activities in new product design teams. In *Proceedings of the ASME Design Theory and Methods Conference*. 351–360.

Rebecca R. Springmeyer, Meera M. Blattner, and Nelson L. Max. 1992. A characterization of the scientific data analysis process. In *Proceedings of the 3rd IEEE Conference on Visualization*. 235–242. Available at http://dl.acm.org/citation.cfm?id=949685.949728.

STEM to STEAM. 2013. What Is STEAM. Retrieved May 22, 2014, from http://stemtosteam.org/.

Robert J. Sternberg. 1999. *Handbook of Creativity*. Cambridge University Press. Available at http://books.google.com/books?id=d1KTEQpQ6vsC.

Jaime Teevan, Edward Cutrell, Danyel Fisher, Steven M. Drucker, Gonzalo Ramos, Paul André, and Chang Hu. 2009. Visual snippets: Summarizing Web pages for search and revisitation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*. ACM, New York, NY, 2023–2032. DOI:http://dx.doi.org/10.1145/1518701.1519008

Elizabeth Thiry, Sian Lindley, Richard Banks, and Tim Regan. 2013. Authoring personal histories: Exploring the timeline as a framework for meaning making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*. ACM, New York, NY, 1619–1628,

James J. Thomas and Kristin A. Cook. 2005. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society Press. Available at http://books.google.co.uk/books?id=DybZPAAACAAJ.

James J. Thomas and Kristin A. Cook. 2006. A visual analytics agenda. *IEEE Computer Graphics and Applications* 26, 1 (January 2006), 10–13. DOI:http://dx.doi.org/10.1109/MCG.2006.5

Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. 2012. The Bohemian Bookshelf: Supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1461–1470. DOI:http://dx.doi.org/10.1145/2207676.2208607

Priyamvada Tripathi and Winslow Burleson. 2012. Predicting creativity in the wild: Experience sample and sociometric modeling of teams. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW'12)*. ACM, New York, NY, 1203–1212. DOI:http://dx.doi.org/10.1145/2145204.2145386.

Edward R. Tufte. 1990. *Envisioning Information*. Graphics Press, Cheshire, CT.

Garrick Utley. 1997. The shrinking of foreign news: From broadcast to narrowcast. *Foreign Affairs* 76, 2, 2–10. Available at http://www.jstor.org/stable/20047932.

Velcro. 2013. Velcro Industries History and George de Mestral. Retrieved May 22, 2014, from http://www.velcro.com/About-Us/History.aspx.

Vitruvius. 1914. *The Ten Books on Architecture*. Harvard University Press. Available at http://www.gutenberg.org/files/20239/20239-h/29239-h.htm.

Edward Vul and Harold Pashler. 2007. Incubation benefits only after people have been misdirected. *Memory and Cognition* 35, 4, 701–710. DOI:http://dx.doi.org/10.3758/BF03193308 10.3758/BF03193308.

Andrew Webb and Andruid Kerne. 2008. The in-context slider: A fluid interface component for visualization and adjustment of values while authoring. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '08)*. ACM, New York, NY, 91–99. DOI:http://dx.doi.org/10.1145/1385569.1385586

Andrew M. Webb and Andruid Kerne. 2011. Integrating implicit structure visualization with authoring promotes ideation. In *Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries (JCDL'11)*. ACM, New York, NY, 203–212. DOI:http://dx.doi.org/10.1145/1998076.1998116

Andrew M. Webb, Rhema Linder, Andruid Kerne, Nic Lupfer, Yin Qu, Bryant Poffenberger, and Colton Revia. 2013. Promoting reflection and interpretation in education: Curating rich bookmarks as information composition. In *Proceedings of the 9th ACM Conference on Creativity and Cognition*. 53–62.

Mark Weiser and John Seely Brown. 1997. The coming age of calm technology. In *Beyond Calculation*. Copernicus, New York, NY, 75–85. Available at http://dl.acm.org/citation.cfm?id=504928.504934.

Ryen W. White, Bill Kules, Steven M. Drucker, and M. C. Schraefel. 2006. Supporting exploratory search: Introduction. *Communications of the ACM* 49, 4, (April 2006), 36–39. DOI:http://dx.doi.org/10.1145/1121949.1121978

Merryl J. Wilkenfeld and Thomas B. Ward. 2001. Similarity and emergence in conceptual combination. *Journal of Memory and Language* 45, 21–38.

Max L. Wilson, Paul Resnick, David Coyle, and Ed H. Chi. 2013. RepliCHI: The workshop. In *Proceedings of the CHI'13 Extended Abstracts on Human Factors in Computing Systems (CHI EA'13). ACM, New York, NY, 3159–3162.*

Allison Woodruff, Andrew Faulring, Ruth Rosenholtz, Julie Morrsion, and Peter Pirolli. 2001. Using thumbnails to search the Web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01)*. ACM, New York, NY, 198–205. DOI:http://dx.doi.org/10.1145/365024.365098

William Wright, David Schroh, Pascale Proulx, Alex Skaburskis, and Brian Cort. 2006. The sandbox for analysis: Concepts and methods. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*. ACM, New York, NY, 801–810. DOI:http://dx.doi.org/10.1145/1124772.1124890

Lauren Yapalater. 2013. The 22 Most Fabulous Beyoncé Moments from the Inauguration. *BuzzFeed.* Retrieved May 22, 2014, from http://www.buzzfeed.com/lyapalater/the-most-fabulous-beyonce-moments-at-the-inaugurat.

Lauren Yapalater and Katie Notopoulos. 2013. 23 Reasons Sasha and Malia Stole the Inauguration. *BuzzFeed.* Retrieved May 22, 2014, from http://www.buzzfeed.com/lyapalater/23-reasons-sasha-and-malia-stole-th e-inauguration.

Ka-Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. 2003. Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, 401–408. DOI:http://dx.doi.org/10.1145/642611.642681

Received August 2012; revised February 2014; accepted February 2014