Search Support for Exploratory Writing

Jonas Oppenlaender[1(✉)], Elina Kuosmanen[1], Jorge Goncalves[2],
and Simo Hosio[1]

[1] University of Oulu, Oulu, Finland
{jonas.oppenlaender,elina.kuosmanen,simo.hosio}@oulu.fi
[2] University of Melbourne, Melbourne, Australia
jorge.goncalves@unimelb.edu.au

Abstract. Writing articles involves searching, exploring, evaluating and reflecting upon different perspectives. To this end, online search engines are commonly used tools to support writing. However, online search engines, such as Google, fall short in supporting complex queries that satisfy multiple criteria simultaneously. In this paper, we present our studies with *GAS*, a crowd-powered tool that allows writers to discover viewpoints, solutions and ideas that best fulfil multiple criteria simultaneously. Our user studies validate GAS as a beneficial companion to online search engines in supporting writing. We found that GAS helps people come up with ideas and write with more confidence, resulting in a higher self-reported article quality and accuracy when compared to only using an online search engine. Through our experiments, we also develop an understanding of the distinct process that people employ when searching for and exploring open-ended, subjective information to support exploratory writing.

Keywords: Exploratory writing · Crowdsourcing ·
Exploratory search · Complex search · Creativity support ·
Qualitative insights

1 Introduction

Writing an article typically requires creative thinking and foraging for further information [37]. *Exploratory writing* [6,29] has been used in the literature to describe the cognitive exploration of information with the aim of gathering information for a writing piece, primarily in the context of fictional writing or teaching. While this term has no clear and widely agreed upon definition in academic literature, in this paper we define exploratory writing as writing aimed at exploring, understanding and reflecting upon different perspectives of a specific topic. To this end, writers can reach vast amounts of information effortlessly through online search engines (OSE). These search engines are engineered to support the most frequently occurring search queries with one search dimension at a time (*e.g.*, cheap, best, fastest) [12,34]. However, online search engines fall short

in supporting *complex* search queries which reflect a personal multidimensional information need [1,7].

As an example, consider the task of writing an article about movies that simultaneously have the *worst* actors and the *greatest* visual effects. You will, most likely, discover search results about one of those subjective criteria at a time, but not results that best fulfil both requirements. This is the case for two main reasons. First, webmasters typically optimise the content of their websites for the most popular information needs, a practice called Search Engine Optimisation (SEO). Second, online search engines also optimise their operations to support the most popular search queries. Yet, there is an identified need for supporting multidimensional search and discovery of unpopular information [35]. For instance, people suffering from back pain may want to discover *fast, cheap and long-lasting ways of treating back pain* – a real-world information need for which online search engines are not well-equipped [21].

In this paper, we present our studies with *GAS*, an adapted version of the decision support concept presented by Hosio et al. in [19]. GAS allows authors to explore and draw inspiration from a knowledge base of open-ended, subjective data on a specific topic. Thus, it aims to support exploratory writers in both searching and exploring information. In our primarily qualitative inquiry, we investigate GAS in supporting exploratory writing. More specifically, we evaluate how it supports complex search tasks when writing a short article. The results from our studies validate GAS as a useful companion to major online search engines in discovering information. GAS helps writers in developing an idea for an article more quickly than online search engines. GAS also reduces the time required to explore the space of possible solutions. Finally, according to self-reported assessment, GAS helps people write articles of higher quality and with higher confidence. Through participant observation and qualitative insights, we also contribute a writer-oriented understanding [22] of the distinct strategies and composing processes that people adopt when searching for open-ended, subjective information in supporting exploratory writing. Finally, we discuss our results in light of related work and focus on the challenges that people encounter when researching a topic online for exploratory writing.

2   Related Work

Our work is situated at the intersection of online tools for supporting writing, information discovery, and creativity. Before delving into these topics, we elaborate on exploratory writing to provide context for our work.

2.1   Exploratory Writing and Complex Search

Exploratory writing is a term with no widely agreed upon academic definition. Lee et al. articulated exploratory writing as "the act of pre-writing for a final writing piece to examine a writer's own thoughts and beliefs and discover more about his/her inner mind and connect to the outside world" [29]. Pre-writing is a

"stage of discovery in the writing process when a person assimilates his 'subject' to himself" [39]. Exploratory writing, from this perspective, is "thinking-on-paper [...] to discover, develop and clarify our own ideas" [6]. It is a preparatory activity that is conducted to gather topical and non-topical material in support of the topical progression of a final writing piece [28]. Murray referred to this activity as "unfinished writing" [31].

Exploratory writing is motivated by a gap in the knowledge of the writer. Exploratory writing may be used to focus ideas and investigate questions [6]. Exploratory writing thus involves and develops critical thinking and reflection [29]. Consequently, one domain in which mentions of the term exploratory writing can be found in the literature is teaching of writing and active learning, *e.g.* in [6,17,29]. Bean, for instance, recommends exploratory writing as an informal writing assignment for use in classrooms to promote critical thinking and learning in students [6]. Haushalter used exploratory, informal writing as a teaching method in an introductory Biochemistry class to engage students in critical thinking and as a mechanism for students to engage with domain-specific knowledge [17]. In this context, exploratory writing may be regarded as exploring and discovering through language [31].

Writers engaged in exploratory writing strive to get an overview over a topic by understanding different perspectives. The task of exploratory writing is an ill-structured problem [44] that requires accessing a large amount of potentially relevant information. Exploratory writing therefore, by definition, involves searching for new information. As such, exploratory writing is a knowledge-crystallisation activity [37] and documents the outcome of an *exploratory search*, with the piece of writing as its final output. Exploratory search refers to the use of "a combination of searching and browsing behaviours to navigate through (and to) information" [49]. Exploratory writers engage in this exploratory search [49] in the context of an information-seeking problem that is open-ended and multifaceted. More specifically, exploratory writers engage in *complex search* tasks. A complex search involves complex information [1] and decisions on how to assess the obtained search results [7]. Such complex search tasks require a disproportionate amount of attention [16], because complex information needs are different from simple information needs, where users search for single information elements [1]. As exploratory writing requires reasoning over multiple pieces of evidence [47] and integrating multiple elements of information, the involved complex search tasks incur a high cognitive load on the writer [46].

Contrary to the literature, we see exploratory writing as *final* writing aimed at exploring, understanding and reflecting upon different perspectives on a non-fictional topic. This type of writing is conducted to stimulate discourse about a topic, investigate questions, eliminate alternatives, or contrast different concepts in written form [6]. In this writing activity, exploratory writers must consider a trade-off between two competing aspects in their search strategy. First, the aim of the writer is to search and find the required information quickly to support the ongoing writing effort. Second, the writer may want to browse and explore the information space to either find the optimal answer or to contrast answers.

To this end, *exploratory search systems* "provide guidance in exploring unfamiliar information landscapes" [49], and therefore support exploratory writing.

The design space for exploratory activities is clearly emerging [3,16,26,27]. For instance, Kittur et al. investigated the costs and benefits of structuring information during the sensemaking process [26]. In this paper, we provide the first systematic investigation of exploratory writing from the perspective of exploratory search systems and complex search for online information. We contribute to shaping the space of exploratory writing in the extracurricular context of non-fictional writing. In the following section, we review key online tools and resources that support exploratory writing and that are most closely related to our work.

Table 1. Online websites and their affordances in supporting exploratory writing.

Website type	Example	Exemplary affordance
Online search engines	Google	Information on any topic, primarily in unstructured form
Open knowledge bases	Wikipedia	Basic information in unstructured form (with exception of tables)
Structured knowledge bases	Wikidata	Structured data, primarily created for consumption by machines
Social media	Twitter	Short information about trending topics and events
Question & answer websites	Quora	Asynchronous answers on any topic
Crowdsourcing platforms	Mechanical Turk	Scalable information from anonymous contributors
Personal social networks	Facebook	Personalised opinions and recommendations

Fig. 1. Example of a complex search task as posted on the AskReddit community.

2.2 Online Support for Writing

Online search engines, such as *Google* and *Bing*, are common tools for discovering, exploring, understanding and reflecting upon different perspectives related

to a given topic. Prior work has highlighted the need to support finding unpopular information [48]. The lookup-based information retrieval model of online search engines is, however, best suited for fact-finding and question-answering with focus on one key search criterion [34,49] and it is not feasible for online search engines to cover the "long tail" of information needs [10]. Instead, online search engines are aligned with their users' search behaviour which follows a power law distribution [35].

Searching for information on the Web is not limited to online search engines. Table 1 gives an overview of the affordances of a selected, but not exhaustive, list of websites that writers may consult in support of their exploratory writing. For instance, open knowledge bases, such as Wikipedia, provide information on a broad spectrum of subject matters. The extended social network of the exploratory writer may further offer support in the exploration of a topic [36].

Relating to the example given in the introduction, Fig. 1 depicts an AskReddit[1] post in which a Reddit user inquires about *movies that are average until a twist made them epic.* The high number of replies in this discussion (over 4,700 comments in less than 19 h) demonstrates the difficulties of getting a comprehensive overview of the answers provided in online communities. While the rich discussion answers the question, such information requests support the *planning* of a writing piece, but not the immediate *writing.* While social computing and social media websites, such as Reddit, Twitter and Quora, can provide answers to complex queries, results are provided asynchronously and not in a timely fashion, therefore thwarting the writing effort.

Ultimately, organising and curating user-created content is a challenging task, and the need for case-specific approaches for acquiring and organising information is clear. Crowdsourcing is one contemporary approach to collect vast amounts of information cost-effectively and accurately.

2.3   Crowd-Powered Search and Writing Tools

The potential of crowdsourcing in supporting writing has been established in the literature [9]. Crowdsourcing has further been applied to searching for information – a key activity in supporting writing. For instance, Bernstein et al. augment a search engine with crowdsourced results for uncommon search queries [10]. In their DataSift tool [34], Parameswaran et al. ask the crowd to provide solutions for queries that cannot be answered by the current online search engines. These queries include multiple dimensions, human judgement calls, and ambiguity. CrowdDB [15] is another example of a tool addressing queries that search engines fail to answer. The authors discuss the difficulty of processing crowd-contributed data that does not follow the closed-world assumption of traditional database systems. CrowdQ [13] uses the crowd to decompose complex queries into query patterns that facilitate answering all instances of future complex queries. In their CrowdSearcher platform [11], Brambilla et al. take a community-based crowdsourcing approach, thus leveraging the diversity of communities with different

---

[1] www.reddit.com/r/AskReddit/.

characteristics and capabilities. With Aardvark [18], a web-based social search engine, users were able to ask subject matter experts in their extended social network for information. Witkeys in China are another example of knowledge market systems in which users can pose questions that cannot be answered easily by online search engines [50].

While the above crowdsourcing systems are able to quickly produce an answer to a complex search task, they do not inherently support the exploration of a multidimensional information space. Questions posed in these systems are one-off queries, and the results may arrive much later. Exploring a topic from multiple perspectives requires multiple queries, possibly in a cascaded fashion [13], which makes the search expensive and time-intensive. Furthermore, the space of possible solutions cannot be explored after the information was retrieved without the user expending cognitive effort on analysing and comparing the search results. There is a need for supporting the exploration of the multidimensional information space in a more swift and effortless way.

To this end, Hosio et al. created a crowd-powered personal decision support system that can be used to collect a knowledge base on arbitrary topics [19,20]. Such knowledge bases can be used to discover best-matching solutions, defined by a set of optimal criteria values, to an arbitrary problem. In this paper, we adopt the approach of Hosio et al. for information discovery. We investigate our tool's ability for supporting exploratory writing instead of decision-making.

3   System Description

To support exploratory writing, we obtained and used *AnswerBot*, the crowd-sourced decision support system presented in by Hosio et al. in [19]. The system provides decision support in form of short textual ideas that can be sorted according to different criteria (see Fig. 2). A crucial consideration in this system, and by extension in GAS, is the data model that facilitates the information discovery. The data model structures a problem space into (1) a set of *ideas* to address the problem, (2) a set of *criteria* (search dimensions) to consider in evaluating the ideas, and (3) the assessment of how each of the ideas fulfil the criteria. Both the ideas, the criteria, and the rating of each idea across all criteria are provided by the crowd.

The content of the knowledge base can be collected from any source that involves the crowd, *e.g.* online labour markets, such as Amazon Mechanical Turk, or any online platforms by posting a link to the data collection mechanism. In discovering the best-matching ideas, the system calculates the closest-matching ideas based on the Euclidean distance of crowdsourced ground-truth values to the desired criteria configuration. In doing so, the system applies *wisdom of the crowd* [45]. We direct the reader to the original article on the system and the newer account for a detailed overview [19,21].

We made minor modifications to AnswerBot's visual query interface by removing the branding and the links to the homepage. We named our system *GAS* – a lightweight, web-based tool for structured exploration of crowdsourced
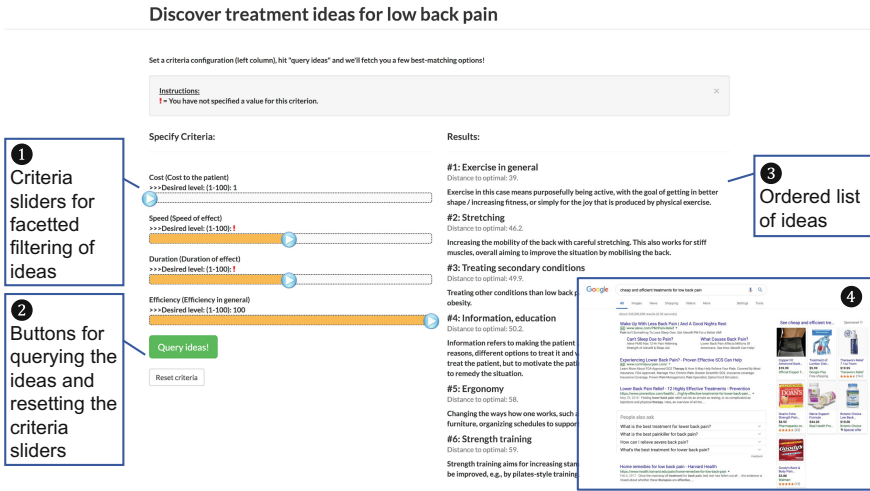
Fig. 2. Main image: searching for *cheap* and *efficient* treatments for low back pain with GAS. Bottom right (❹): results for the same query in Google (search location in this example: Washington, D.C.).

knowledge, to support writing. By design, the system is intended for querying and discovering crowdsourced, open-ended ideas. GAS is not a replacement for Web search engines, but aimed at helping writers in exploring the information space as a complementary tool.

Users of the GAS system set the desired criteria combination by using sliders in the left column of the GAS look-up interface (see ❶ in Fig. 2). Corresponding results are retrieved (with a click on a button ❷) and displayed in the right column (❸ in Fig. 2). Figure 2 compares a search on GAS and Google for *cheap but efficient treatments for low back pain*, as an example. This exact case with the original system was discussed in-depth in [21].

4 Evaluation

To investigate to what extent GAS can support exploratory writing, we set up two distinct studies. Next, we provide details of the knowledge bases used in the studies.

4.1 Crowdsourced Knowledge Bases

The knowledge bases in AnswerBot are collected dynamically over time to answer topic-specific information needs: they are not transferable between different problem domains. For this paper, we obtained and reused three knowledge bases collected by the AnswerBot platform and documented in [8,21]. The three data collection campaigns are summarised in Table 2. Data for T1 and T2 was collected "organically" in a several months long campaign promoted on social media

Table 2. Overview of the data collection campaigns.

	T1. RACISM	T2. LBP	T3. DIETS
	What are good means to prevent everyday racism?	What are good ways to treat low back pain?	What are good weight-loss diets?
Contributors	1,515	353	70
Ideas	84	108	21
Criteria	5	4	6
Assessments	13,654	9,001	8,607
Collection	Organic	Organic	Prolific
Criteria source	Sociologists	MDs, physiotherapists	Website visitors

outlets, forums and mailing lists. T3 was collected via Prolific[2], a crowdsourcing platform.

4.2 Study 1: Writing with and Without GAS

We conducted a within-subject study with 24 participants (15 male, 9 female, age range 21 to 38 years, $Mean = 27.2$ years, $SD = 4.36$ years, $Mdn = 25$ years). Participants were recruited through a mailing list and posters on our campus. Participants included 3 recently graduated Master students, 1 Associate Professor, 1 MBA, 2 bachelor students, 3 PhD students and 14 Master students. Given the participants' academic background, we consider them as knowledgeable in exploring topics and writing essays. Master students at our institution are regularly engaged as research assistants, which includes researching information and writing. All participants reported being familiar with researching information online, but not having experience with complex searches. Only two participants reported using logical keywords and operators in their daily searches. Participation in the study required approximately 30 min and was rewarded with a cinema ticket voucher worth US $12.

The participants conducted the study independently on a computer provided by us, situated in a semi-controlled work environment. First, a guided tutorial consisting of written instructions and oral explanations provided by a researcher familiarised the participant with the interface of GAS, using topic T3. The writing part of the study consisted of two preset writing tasks (T1 and T2), an online survey, and a follow-up interview. Similar preset essay tasks have been used earlier, *e.g.* in investigating computer-aided creative writing [41].

In the writing part, each participant was asked to write two short articles. Participants were intentionally given a tight time limit of only 5 min for each writing task to test the participants' ability of finding and exploring information quickly. The available criteria in the "mitigating racism" task (T1) were

---

[2] www.prolific.ac.

as follows: (1) ability to influence, (2) speed of solution, (3) cost for society, (4) efficiency, and (5) legality of solution. For the "curing low back pain" task (T2), the criteria were: (1) cost to the patient, (2) speed of effect, (3) duration of effect, and (4) efficiency. Each of the writing tasks consisted of one topic (T1 or T2) and a randomly selected combination of two criteria values (see Fig. 3). The tasks were completed under one of two conditions:

– `OSE`: supported only by a major online search engine, such as Google, and
– `GAS+OSE`: supported by the combination of GAS and a major online search engine, such as Google, *i.e.*, participants were allowed to use both the search engine and GAS, on their own volition.

Our study design, on purpose, did not include a choice of using only GAS, as this would force the users to use a non-realistic and somehow enforced setup and thus decrease the ecological validity of the results. Instead, the participants had a choice of using the tools as they saw fit. The order of the conditions and topics was counterbalanced to avoid carry-over and learning effects. Each task was completed in an instance of Etherpad[3], pre-filled with task instructions.

Fig. 3. Participant P1's tasks and articles.

After completing the writing tasks, we collected subjective responses, reported on a 7-point Likert scale, to questions relating to five aspects of the writing task:

– mental effort (*e.g.*, for searching, remembering, thinking, and deciding) from 1 (Low) to 7 (High),
– difficulty of finding the initial idea for the article, from 1 (Very Difficult) to 7 (Very Easy),
– difficulty of finding content for the article, from 1 (Very Difficult) to 7 (Very Easy),
– task difficulty, from 1 (Very Easy) to 7 (Very Difficult),
– perception of usefulness of the information provided by each tool, from 1 (Not At All Useful) to 7 (Extremely Useful),
– how well the tools supported the writing, from 1 (Not At All Helpful) to 7 (Extremely Helpful),

---

[3] www.etherpad.org.

– self-assessed adherence to the task criteria, from 1 (Not At All) to 7 (To A Great Extent),
– self-assessed quality of the written articles, from 1 (Very Poor) to 7 (Excellent),

and the overall preference (OSE versus GAS+OSE). For a lightweight evaluation of mental effort and task difficulty, we adopted the questions from Siangliulue et al. [43].

Finally, we conducted one-on-one interviews with each participant immediately after the questionnaire was completed. The interviews were semi-structured in nature and focused on further exploring the choices of the participant and the observed behaviour. After the study ended, we analysed the interview data using thematic coding [38]. Thematic coding is an approach to identify themes in qualitative data. We applied this approach to construct a framework that guided our investigation.

4.3   Study 2: Writers' Workshop

To gain more insights into the writing and search strategies, as well as to validate the strategies that we derived from the interview data, we conducted a concluding workshop with six participants (W1-W6, Bachelor and Master level students, two of which had participated in the main study), recruited from the same University campus. The workshop was led by two researchers, one of which acted as scribe. In the workshop, one participant volunteered to think-aloud and give verbal instructions to one of the researchers who operated the GAS system. After 15 min of searching and exploring for information concerning both topics T1 and T2, the other workshop participants were allowed to join the process and discuss freely around GAS and its functionalities. We include findings from the workshop in the qualitative data analysis and in the discussion.

5   Results

While our investigation is primarily of qualitative nature, we also report quantitative insights in this section.

5.1   Quantitative Data Analysis

In the following, we report statistics from the analysis of the Likert-scale questionnaire responses (see Fig. 4). We compare the two writing conditions (OSE and GAS+OSE) using Wilcoxon's signed rank test with continuity correction. We used an alpha level of .05 for all statistical tests.

Finding an initial idea for the article was rated significantly easier when using GAS as a complementary tool, $Mdn = 5$, compared to using OSE, $Mdn = 4$, $V = 40.5$, $p = .008$. Participants also responded that finding content for the article was easier when using GAS, $Mdn = 6$, compared to the OSE condition,
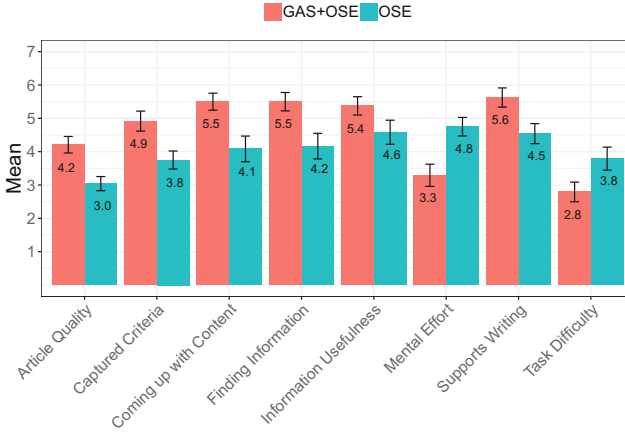
Fig. 4. Mean Likert-scale responses as reported by the participants.

$Mdn = 4$, $V = 64.5$, $p = .013$. Overall, participants thought that GAS supported their writing better, $Mdn = 6$, than the information provided by the online search engine alone, $Mdn = 5$, $V = 30.5$, $p = .005$). Using GAS also significantly reduced the participants' mental effort, $Mdn = 3$, compared to OSE, $Mdn = 5$, $V = 254$, $p = .001$. However, the data afforded by GAS was not deemed to be significantly more useful in solving the complex search task, $Mdn = 6$, as compared to the OSE condition, $Mdn = 5$, $V = 36$, $p = .015$). Nevertheless, the task was perceived to be less difficult under condition GAS+OSE, $Mdn = 2$, compared to condition OSE, $Mdn = 3.5$, $V = 178.5$, $p = .046$. An overwhelmingly high percentage of the participants (87.5%) preferred to solve the task with a combination of GAS and the search engine, 12.5% preferred to only work with Google. None of the participants preferred using the data solely provided by GAS.

The number of words in the articles did not significantly differ between the OSE ($Min = 12$, $Max = 185$, $Mean = 82.2$, $Mdn = 80$) and GAS+OSE ($Min = 26$, $Max = 665$, $Mean = 102.8$, $Mdn = 74.5$) conditions ($V = 137.5$, $p = .5$), even though the texts written with GAS had on average a higher number of words. Figure 3 depicts the two articles written by participant P1. Both articles address the criteria of the complex task to some extent. We note, however, that the article written under condition GAS+OSE is more nuanced. This intuition is confirmed by the self-reported ratings of the participants. Overall, the participants thought that their own articles written with support of the information provided by GAS, $Mdn = 5$, were better matching the task criteria than with the support of the search engine, $Mdn = 3.5$, $V = 50.5$, $p = .003$. The participants rated the quality of their own article written with support of GAS significantly higher, $Mdn = 4.5$, than the one written with help of the online search engine, $Mdn = 3$, $V = 17$, $p < .001$. When asked to compare both articles, 14 participants (58.33%) said that the article written with both GAS and

the online search engine was better. Five participants rated the article written with support of the online search engine as better, and another five participants said the two articles were about the same quality.

5.2 Qualitative Insights into Search and Exploration Strategies for Exploratory Writing

Based on the thematic grouping of the coded interview data, we developed a framework that is indicative of how participants engaging in exploratory writing in our study addressed their information needs (see Fig. 5). While our framework follows a stage-process model of writing, as defined for instance by Rohman [39], we acknowledge that writing is an iterative cognitive activity [14] and that there may be iterations between searching, idea generation, refining and writing. Given how we derived the model, the stages of our framework are reflected in the headings of our following qualitative analysis.

Fig. 5. The main activities and influencing factors observed in exploratory writing

Our framework relates to many existing models for information discovery and writing. For instance, Kerne and Smith's information discovery framework "examines interactions of human cognition with digital information in tasks that draw upon information resources while developing ideas" [25]. This framework is concerned with inner knowledge states and knowledge transformations, whereas our framework is more focused on the immediate writing process and the external factors that influence the exploratory search. Compared to Marchionini's exploratory search activities [30], our model is concerned with search, idea generation and writing. Lastly, Flower and Hayes's writing model [14] is a detailed model of the writing process, but was written before the inception of the Internet. For an overview of related models from the context of information retrieval and exploratory search, we direct the reader to White and Roth's synthesis lecture on exploratory search [49].

5.3   Challenges in Writing with Google Alone

The topics were perceived as challenging and the combination of the criteria seemed contradicting at first (P13). It does not come as a surprise that combining the two criteria in an article was perceived as *"most difficult"* (P4).

> *"The challenge was writing something useful. All I wrote was pieces and bits from topics. It's not a good article."* (P24)

The time limit was considered a major challenge by several participants. During the study, we noticed that participants had to click away cookie consent overlays, and several participants ended up in dead ends due to being mislead about the contents of websites or information being locked behind paywalls. Integrating information from disparate sources (P14) and finding reliable information (P11) were further reported as challenges. Advertisements on the search engine (*c.f.* Fig. 2), in particular, were not considered trustworthy by the participants. For some participants, this distrust for advertisement extended to the embedded answers provided by Google. As W5 noted, *"embedded answers oversimplify what I need. Google probably finds simple stuff that it can easily parse to promote as the embedded answer."*

5.4   Searching for Information Without and with GAS

In this section, we describe how the participants tackled the complex search task under the OSE condition and how GAS supported this task.

Searching Without Support of GAS. Participants usually turn to Google as their main search engine and draw on their previous experience and prior knowledge.

> *"Usually before writing, I try to think what knowledge I already have about the topic, and based on that I start searching the information, usually using Google."* (P1)

If no prior knowledge about a topic exists, the participants turn to Wikipedia for basic information or consult peers and domain experts.

We observed that some participants followed a two-stage query strategy in solving the complex search task. In this *criteria-driven* approach, participants used the criteria to drive results, first searching for one criterion on the search engine, then filtering among the results for the second criterion.

> *"First, I searched for one criteria. Then from those results, I figured out the other criteria."* (P4)

P3 reported that he *"read at least five articles on Google, found similarities among the results, and wrote based on that."* This approach was however limited by the insufficient quality of the results provided by the search engine, which lead P21 to comment: *"What I found was too broad. I could not match the criteria."* Some participants perceived the online search engine not useful in solving the complex search task.

"I doubted that the search engine would be able to produce results, so I did not even try." (P1)

The search engine results were described as *"soup of information"* (P5), providing *"results optimised for only the most searched things"* (P14), *"links to general articles"* (P19), and a *"confusing amount of information"* (P3). *"Some of the search engine results were hard to read,"* noted participant P1 who skimmed through legal texts. Some participants expected to see embedded answers (*e.g.* P5), but did not receive such information in the search engine results.

Searching, Supported by GAS. GAS was conversely perceived as *"specific and goal-oriented"* (P15) and *"straight forward"* (P18) in retrieving information for a specific topic. GAS allows to *"focus your search instead of browsing"* (P23). The *"condensed and easily readable data format"* (P1) of GAS was perceived as useful in solving the task. GAS was reported as making the task *"easier"* (*e.g.* P13). Since the results provided by GAS are organised around a topic, the tool *"provides a headstart into the article"* (P24), affording a *"kind of intuitiveness"* (P23).

"On Google, you need to think about the search phrase. [GAS] simplifies finding answers based on several criteria." (P5)

Participants were able to find information with GAS faster than with the online search engine. P1, for instance, noted that she *"did not have to make several searches."*

"While the search engine provides results for only one criteria, [GAS] weighs in all criteria." (P13)

GAS directly provides relevant content, without having to go through links, like on the online search engine. This allowed the participants to compare findings of GAS and the online search engine, if time permitted. Participants used GAS for getting an overview of the information space and helping to *"understand answers in a larger context"* (P5). The structured data afforded by GAS enables the writer to compare results, explore different viewpoints, and systematically consider the trade-offs between answers to a question. GAS affords a ludic exploration of the information space spanned by the available ideas and criteria.

Status Quo Bias. We noticed a status quo bias [40] among certain participants. Google was reported by these participants as being the *"easiest way"* (P20) of retrieving information. The participants from this group did not try GAS and simply completed the tasks using Google.

"If you have a tool and it's very useful, and you already use it before, you go directly to use it instead of the new thing." (P20)

Another participant from this group said that he would *"need to practice [GAS] before using it"* (P16). This answer surprised us, given that we made sure to give each participant an exhaustive interactive introduction to GAS.

In the group of users who preferred the search engine, interacting with the search engine was perceived as being more natural than interacting with the sliders in the user interface of GAS.

> *"I had to read through all the criteria descriptions in [GAS]. When using Google, the question is entered and answers are given."* (P22)

Overall, a clear majority of 21 out of 24 participants preferred to use the combination of GAS and Google in solving the complex writing tasks. As P5 remarked, *"the combined approach would produce best results."*

5.5   Inception of the Initial Idea

We observed four different approaches relating to how the information afforded by GAS informed the participants' search strategies [5] and search tactics [4] for finding the fundamental idea of the article. We call these approaches source-of-inspiration, knowledge-driven, structure-driven and hybrid.

Source-of-Inspiration Approach. The information provided by GAS was deemed insufficient for longer articles, but good for inspiration. The majority of our participants used GAS as the sole source of inspiration for finding the main idea for their article. Some of the participants used the answers provided by GAS as-is, copying and pasting the answers *"without thinking"* (P17). GAS is *"for getting started,"* noted P18. For more inquisitive users, GAS *"provided a list of actionable keywords to search for in the search engine"* (P13). Another approach under this theme was to summarise existing answers from GAS.

Knowledge-Driven Approach. Participants who used this approach used their intuition and prior knowledge for coming up with an idea for the article and then *"continued that direction"* (P2), with or without support of GAS or the search engine. This approach was particularly prevalent in the group of participants affected by the status quo bias. P4 noted that she *"could not find anything matching on Google, only a general idea"* and that she *"used [her] general knowledge to solve the task."* Past research has shown that emerging candidate solutions may guide the user interaction behaviour [49]. For some participants, this internal guidance was so strong that they did not make use of the search engine or GAS to their full extent (or even at all). Digging deeper into this behaviour, these participants said they knew the search engine would not have been able to solve their information need. Instead, they decided to formulate their initial intuition, informed by prior knowledge, into written form.

Structure-Driven Approach. In this approach – a subset of the knowledge-driven approach – participants first came up with an idea and structure for the article on their own, and then turned to GAS and the search engine. P5 discussed his writing process in detail, which consisted of first understanding

the requirements, writing headings for each section, and finally finding pieces of information that support each heading. P14 also first sketched out a rough idea before turning to the search engine. The search engine was used to fill in the blanks in the structure. GAS conversely was perceived helpful in coming up with the initial structure for the article.

Hybrid Approach. Many participants combined GAS and the search engine, using GAS to get a *"rough idea about a question and possible directions to look at"* (P14). P5 reported that he *"started with GAS, choosing criteria. From there, [he] expanded to understand more on Google about these answers and to develop a broader understanding."*

5.6   Idea Refinement

A clear majority of the participants turned to the search engine as an external source for refining the article, primarily with the intention of finding more information. We noticed a *berrypicking approach* [5] for integrating information. Participants skimmed through the snippets provided by the search engine and GAS, and harvested elements to integrate into the writing. P2 reported that he *"searched for articles [on the search engine], read through the information, and integrated relevant pieces into the article."* With support of GAS, this approach was more profound and all-consuming. According to our observations, significantly more information was copied and pasted when using GAS.

Another way of refining the article was through self-reflection, *i.e.*, not using external sources, to bring the article into one's "personal context" [39]. As P1 reported, he had to rewrite the text to *"sound more like me."* P21 describes his experience in refining the initial idea for the article: *"I tried to improvise and add more. I tried to make it more related to the criteria."* Further activities included reading through the whole article for coherence, and pruning of gathered information. We noticed a change in the writing behaviour when using GAS, compared to the OSE condition. With support of the search engine, participants reported that the idea refinement was largely *"writing, trying to find support, then edit and improve"* (P17). With GAS, refinement shifted towards editing, rephrasing, and *"skimming through, making it more legible"* (P17). We noticed significantly less own development in the texts of the participants when using GAS.

5.7   Trusting the Answers Provided by GAS

The crowdsourced knowledge base was accepted by the participants as being trustworthy. One participant compared the knowledge base to Wikipedia, stating that he *"completely trusted GAS. It felt trustworthy"* (P17). When asked why, the participant said *"for some reason I took it for granted. I don't know why"*. Participants considered the content of the crowdsourced knowledge base to be a mix of specialist and non-specialist opinions. This mix was perceived as more trustworthy than the search engine results.

However, several participants brought up the issue of the provenance of the crowdsourced information, due to not being able to confirm the source of the provided information. An interesting outlier to this behaviour was one workshop participant who wished to verify all ideas suggested by GAS by finding articles on diets from *e.g.* the National Health Service (NHS) and Google Scholar. This particular participant (P18) had a background in nutritional sciences, which may partially explain this behaviour given their expertise in the matter: *"In order to use it, I would personally want to check the validity of the data."* While P18 wished to see references, he acknowledged that it is difficult to provide references for people's opinion. In the end, users trusted their common sense when it came to integrating information from the crowd-powered knowledge base into the article.

6   Discussion

The pervasive access to a seemingly unlimited amount of information online has transformed the process of writing. Insights, debate, and argumentation with varying degrees of reliability from a plethora of different sources and on practically any imaginable topic are always just a few Google searches away [5]. Several online tools exist to help navigate this *information overload* for different purposes, some of which we list in Table 1. The major search engines are doing a wonderful job in enabling discovery among all the user-generated content out there, but they are – understandably – far from perfect in supporting a specific application area such as writing. While the results that most strongly speak for this from our studies are overall more qualitative than quantitative in nature, the response to the Likert-scale (as depicted in Fig. 4) also suggests that tools like ours warrant consideration. On the other hand, it is important to note that any new tools competing for attention have a long struggle ahead: The status quo bias we observed among our participants suggests that many potential users will not give new tools even a chance.

In these studies, we set to explore the perceived value of the tool for its users. While it is certainly important to evaluate the quality of the writing in future studies, we scope our investigation solely on the writing process and the users' experience as something that we can confidently provide results and reflections on. We point to similar publications [23,32] that also evaluated their respective system in this manner.

6.1   GAS as a Tool for Supporting Writing and Information Discovery

In this paper, we emphasise exploratory writing as an activity that involves searching and exploring the information space of the topic [49]. To support such writing, we modified the AnswerBot system by Hosio et al. [19]. Our solution, GAS, ranks open-ended content across multiple criteria simultaneously – an approach versatile for provisioning decision support for real-world problems [21]

but also for discovering ideas that best fulfil a set of search criteria. Thus, it affords a straightforward approach to solving *complex search tasks* [1] that the online search engines have been shown to struggle with [34].

The participants of our studies demonstrated clear stages in their exploratory writing processes (see Fig. 5). Turning our eyes to the challenges encountered when using an online search engine in these stages, we are able to isolate the characteristics of GAS that warrant further exploration and provide useful pointers for designers of similar solutions.

In the *Search Stage*, participants expressed issues with non-relevant information and content motivated by commercial interests on Google. As expected, the clutter of advertisements (see Fig. 2) was often mentioned, but also the embedded answers by Google itself were found as troublesome and to provide too simple solutions for sophisticated information needs. GAS, on the other hand, helps to focus on only the relevant subset of information, a characteristic of high quality Human-Information Interaction [1]. While GAS provides a clutter-free view to relevant information, it does not generalise well: one knowledge base only caters for one problem with a set of criteria and answers at a time.

GAS excels in simultaneously providing multiple relevant "seed ideas" to work with in the *Idea Generation Stage*. This was true when exploring only two simultaneous criteria, and the verdict was even more clear when experimenting with three, four or even five criteria simultaneously in the Writers' Workshop. The opinion on GAS was rather unanimous in the Writers' Workshop, with GAS being dubbed as "far superior" to Google in brainstorming for an initial batch of ideas. However, the crowdsourced knowledge bases clearly failed to provide comprehensive enough content on the ideas, and the participants, not only in the workshop but throughout our study, opted to switch to Google to expand and seek validation for the ideas discovered with GAS.

In the *Idea Refinement Phase*, Google was expressed to be a more compelling option than our tool. This is particularly due to the lack of perceived trust in the source of the provided information. Indeed, the origin of information is one of the greatest factors that influence perceived trust, regardless of the content itself [33]. In this regard, GAS may be perceived as a rather ambiguous source of information: what is a crowd, and how can we communicate the provenance of the crowd-powered information to the end-users? People have an intuitive understanding of what is trustworthy, as also exemplified by the workshop participant wishing to find validation for ideas by GAS from NHS or Google Scholar. But eliciting knowledge from a crowd is far more complicated than processing webpages for inclusion in a search engine's index. Reaching to a large pool of verifiable and trustworthy experts is not an easy feat. In retrospect, in our study we could have investigated such trust aspects as well, since the knowledge base we used on low back pain includes data harvested from clinical professionals only [21]. We leave examining these issues for future work.

6.2 Supporting Creativity with the Crowds

Computer-aided programs have been used to teach and support creative writing (*e.g.* [41]), long before the convenience of the now-ubiquitous Internet. In HCI, creativity support has been positioned as one of the grand challenges for scientists to look into [42]. We argue that crowdsourced support for creativity might play an interesting role in this challenge.

GAS, for instance, allows for an intuitive means to tap into the ideas, viewpoints and personal opinions that a crowd has accumulated over time. Therefore, the usefulness of tools like GAS for creativity grows as a function of the number of diverse, knowledgeable and independent individuals participating in the data collection stage [45]. The trade-off to strike lies between costs and utility: collecting a crowdsourced knowledge base to sufficiently cover a specific topic requires a lot of effort, time, and, in most cases, money.

It is not just the access to diverse ideas that matters in creative writing. The authors should optimally be intrinsically motivated in their writing task [2]. In our studies, we did not specifically measure aspects of creativity, but in the Writers' Workshop it was brought up that GAS might be helpful in shedding light on information that a writer might fail to consider if using *e.g.* Google only. It is the criteria, not the given solutions and ideas, that opened up the investigator's mind to consider *unknown unknowns*: the criteria constitute the proverbial lens that a given problem can be examined through. For instance, we might not discover too many surprising diets from a crowd, but when asking what type of criteria people value in diets, we might get a diverse range of aspects that one might fail to otherwise consider in their thought processes.

Finally, creativity is not something that is reserved for the selected few. Gifted novices are almost as good at judging the level of creativity as seasoned expert authors [24]. Thus, a crowdsourced solution like ours could be useful for a wide range of people in exploring how to expand their ideas and enhance their writing results. But, as implied earlier, we do not suggest crowdsourced approaches as competitors to the established online search engines. Search engines organise everyone's content, whereas GAS organises content harvested with a specific topic in mind. This also highlights the importance of correctly choosing the criteria in the beginning. GAS will not provide any clues for missing criteria, potentially severely limiting its usefulness in creativity support. Of course, online search engines are equally limited to the user-generated content online and further optimised for the most popular, relatively simple search tasks [10].

6.3 Limitations of the Study

We acknowledge several limitations in our study. First, while we made every effort to simulate a realistic work environment, the participants' awareness of being studied most likely affected their work quality. Self-reported introspective analysis of the writing quality may further lead to inaccurate reportings [14,22]. However, the positive findings in the self-reports were supported by the consensus

in the qualitative analysis on how our tool complements online search in the various stages of the exploratory writing process.

Further, the study design may have imposed minor carry-over effects. We addressed this in our study by assigning the participants unrelated topics consisting of randomised criteria. Participants were also required to go through the interactive tutorial before starting the study.

Finally, it is not unfair to question the motives of the participants. The information-seeking problem was extrinsically motivated with a cinema ticket as reward. If the problem was intrinsically motivated, the search strategies and tactics may have differed from what we observed. This needs to be explored in future studies.

7   Conclusion

In this paper, we investigated supporting the search for and exploration of information for exploratory writing. While online search engines excel in supporting single criterion look-up of information, they fall short in providing answers to complex queries consisting of combinations of multiple criteria. Our user studies of search behaviours with GAS, a crowd-powered knowledge base of a specific topic, validate GAS as a useful companion to major online search engines, such as Google, especially in supporting idea generation. Further, the study participants self-assessed the articles written with GAS as being of higher quality and better matching the given writing tasks. Overall, both our quantitative and qualitative data constitute compelling evidence of the potential of crowd-powered tools in supporting writing and creativity.

References

1. Albers, M.J.: Human-information interaction with complex information for decision-making. Informatics **2**(2), 4–19 (2015)
2. Amabile, T.M.: Motivation and creativity: effects of motivational orientation on creative writers. J. Pers. Soc. Psychol. **48**(2), 393 (1985)
3. Andolina, S., Klouche, K., Cabral, D., Ruotsalo, T., Jacucci, G.: InspirationWall: supporting idea generation through automatic information exploration. In: Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition (C&C 2015), pp. 103–106. ACM, New York (2015)
4. Bates, M.J.: Information search tactics. J. Am. Soc. Inf. Sci. **30**(4), 205–214 (1979)
5. Bates, M.J.: The design of browsing and berrypicking techniques for the online search interface. Online Rev. **13**(5), 407–424 (1989)
6. Bean, J.C.: Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom. Jossey-Bass, San Francisco (2001)
7. Bell, D.J., Ruthven, I.: Searcher's assessments of task complexity for web searching. In: McDonald, S., Tait, J. (eds.) ECIR 2004. LNCS, vol. 2997, pp. 57–71. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24752-4_5

8. van Berkel, N., Hosio, S., Goncalves, J., Kostakos, V.: Informed diet selection: increasing food literacy through crowdsourcing. In: Proceedings of the Designing Recipes for Digital Food Futures CHI 2018 Workshop (2018)
9. Bernstein, M.S., et al.: Soylent: a word processor with a crowd inside. In: Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology (UIST 2010), pp. 313–322. ACM, New York (2010)
10. Bernstein, M.S., Teevan, J., Dumais, S., Liebling, D., Horvitz, E.: Direct answers for search queries in the long tail. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012), pp. 237–246. ACM, New York (2012)
11. Bozzon, A., Brambilla, M., Ceri, S.: Answering search queries with crowdsearcher. In: Proceedings of the 21st International Conference on World Wide Web (WWW 2012), pp. 1009–1018. ACM, New York (2012)
12. Collins-Thompson, K., Hansen, P., Hauff, C.: Search as learning (dagstuhl seminar 17092). Dagstuhl Rep. **7**(2), 135–162 (2017). http://drops.dagstuhl.de/opus/volltexte/2017/7357
13. Demartini, G., Trushkowsky, B., Kraska, T., Franklin, M.J.: CrowdQ: crowdsourced query understanding. In: Conference on Innovative Data Systems Research (CIDR 2013) (2013)
14. Flower, L., Hayes, J.R.: A cognitive process theory of writing. Coll. Compos. Commun. **32**(4), 365–387 (1981)
15. Franklin, M.J., Kossmann, D., Kraska, T., Ramesh, S., Xin, R.: CrowdDB: answering queries with crowdsourcing. In: Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data (SIGMOD 2011), pp. 61–72. ACM, New York (2011)
16. Hahn, N., Chang, J.C., Kittur, A.: Bento browser: complex mobile search without tabs. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018), pp. 251:1–251:12. ACM, New York (2018)
17. Haushalter, K.A.: Developing critical thinking in introductory biochemistry through exploratory writing in an electronic collaborative learning environment. FASEB J. **21**(5), A297 (2007)
18. Horowitz, D., Kamvar, S.D.: The anatomy of a large-scale social search engine. In: Proceedings of the 19th International Conference on World Wide Web (WWW 2010), pp. 431–440. ACM, New York (2010)
19. Hosio, S., Goncalves, J., Anagnostopoulos, T., Kostakos, V.: Leveraging wisdom of the crowd for decision support. In: Proceedings of the 30th International BCS Human Computer Interaction Conference: Fusion! (HCI 2016). BCS Learning & Development Ltd., Swindon (2016)
20. Hosio, S., Goncalves, J., van Berkel, N., Klakegg, S.: Crowdsourcing situated & subjective knowledge for decision support. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp 2016), pp. 1478–1483. ACM, New York (2016)
21. Hosio, S., et al.: Crowdsourcing treatments for low back pain. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018). ACM, New York (2018)
22. Hyland, K.: Teaching and Researching Writing, 3rd edn. Routledge, New York (2016)
23. Iqbal, S.T., Teevan, J., Liebling, D., Thompson, A.L.: Multitasking with play write, a mobile microproductivity writing tool. In: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, UIST 2018, pp. 411–422. ACM, New York (2018)

24. Kaufman, J.C., Gentile, C.A., Baer, J.: Do gifted student writers and creative writing experts rate creativity the same way? Gifted Child Q. **49**(3), 260–265 (2005)
25. Kerne, A., Smith, S.M.: The information discovery framework. In: Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS 2004), pp. 357–360. ACM, New York (2004)
26. Kittur, A., Peters, A.M., Diriye, A., Telang, T., Bove, M.R.: Costs and benefits of structured information foraging. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2013), pp. 2989–2998. ACM, New York (2013)
27. Klouche, K., Ruotsalo, T., Cabral, D., Andolina, S., Bellucci, A., Jacucci, G.: Designing for exploratory search on touch devices. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015), pp. 4189–4198. ACM, New York (2015)
28. Lautamatti, L.: Observations on the development of the topic in simplified discourse. AFinLAn Vuosikirja **8**(22), 71–104 (1978)
29. Lee, J., Son, J., Settle, Q.: Exploratory writing in student learning. Int. J. Fashion Des. Technol. Educ. **9**(1), 9–15 (2016)
30. Marchionini, G.: Exploratory search: from finding to understanding. Commun. ACM **49**(4), 41–46 (2006)
31. Murray, D.M.: Teach writing as process not product. Leaflet **71**, 11–14 (1972)
32. Nebeling, M., et al.: WearWrite: crowd-assisted writing from smartwatches. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI 2016, pp. 3834–3846. ACM, New York (2016)
33. Nurse, J.R., Rahman, S.S., Creese, S., Goldsmith, M., Lamberts, K.: Information quality and trustworthiness: a topical state-of-the-art review. In: The International Conference on Computer Applications and Network Security (ICCANS 2011), pp. 492–500. IEEE (2011)
34. Parameswaran, A., Teh, M.H., Garcia-Molina, H., Widom, J.: DataSift: a crowd-powered search toolkit. In: Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data (SIGMOD 2014), pp. 885–888. ACM, New York (2014)
35. Petersen, C., Simonsen, J.G., Lioma, C.: Power law distributions in information retrieval. ACM Trans. Inf. Syst. **34**(2), 1–37 (2016)
36. Pirolli, P.: Powers of 10: modeling complex information-seeking systems at multiple scales. Computer **42**(3), 33–40 (2009)
37. Pirolli, P., Card, S.: Information foraging. Psychol. Rev. **106**(4), 643–675 (1999)
38. Robson, C., McCartan, K.: Real World Research, 4th edn. Wiley, Chichester (2016)
39. Rohman, D.G.: Pre-writing the stage of discovery in the writing process. Coll. Compos. Commun. **16**(2), 106–112 (1965)
40. Samuelson, W., Zeckhauser, R.: Status quo bias in decision making. J. Risk Uncertain. **1**(1), 7–59 (1988)
41. Sharples, M.: Cognition, computers and creative writing. The University of Edinburgh, Edinburgh, UK (1984)
42. Shneiderman, B.: Creativity support tools: a grand challenge for HCI researchers. In: Redondo, M., Bravo, C., Ortega, M. (eds.) Engineering the User Interface: From Research to Practice, pp. 1–9. Springer, London (2009). https://doi.org/10.1007/978-1-84800-136-7_1

43. Siangliulue, P., Chan, J., Dow, S.P., Gajos, K.Z.: IdeaHound: improving large-scale collaborative ideation with crowd-powered real-time semantic modeling. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST 2016), pp. 609–624. ACM, New York (2016)
44. Simon, H.A.: The structure of Ill structured problems. Artif. Intell. **4**(3), 181–201 (1973)
45. Surowiecki, J.: The wisdom of crowds. Anchor (2005)
46. Sweller, J.: Cognitive load during problem solving: effects on learning. Cogn. Sci. **12**(2), 257–285 (1988)
47. Talmor, A., Berant, J.: The web as a knowledge-base for answering complex questions. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 641–651. Association for Computational Linguistics (2018)
48. White, R.W., Bilenko, M., Cucerzan, S.: Studying the use of popular destinations to enhance web search interaction. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2007), pp. 159–166. ACM, New York (2007)
49. White, R.W., Roth, R.A.: Exploratory search: beyond the query-response paradigm. In: Marchionini, G. (ed.) Synthesis Lectures on Information Concepts, Retrieval, and Services #3. Morgan & Claypool, New York (2009)
50. Yang, J., Adamic, L.A., Ackerman, M.S.: Crowdsourcing and knowledge sharing: strategic user behavior on Taskcn. In: Proceedings of the 9th ACM Conference on Electronic Commerce (EC 2008), pp. 246–255. ACM, New York (2008)