Article
Collective Intelligence in Design Crowdsourcing

Jonathan Dortheimer [ID]

MTRL Laboratory, Faculty of Architecture and Town Planning, Technion—Israel Institute of Technology, Haifa 3200003, Israel; jonathan@dortheimer.com or dortheimer@campus.technion.ac.il

Abstract: This study investigates how collective intelligence emerges in crowdsourcing for architectural design. Previous studies have revealed that collective intelligence emerges from collaboration and can outperform individual intelligence. As design is a highly collaborative practice, collective intelligence plays a vital role in the design process. In this study, we compare the structure of two architectural design crowdsourcing systems using several methods. The results of the analysis suggest that design crowdsourcing systems can give rise to the following three types of collective intelligence: (1) discussive, which emerges from a conversation between designers and clients; (2) synthetic, which emerges from a parallel and sequential design development; and (3) evaluative, which is based on the wisdom of the crowd in evaluating and selecting designs. The article concludes with recommendations for collaborative design method.

Keywords: collective intelligence; creative crowdsourcing; design methods; crowdsourcing; architectural design

1. Introduction and Background

Crowdsourcing is an umbrella term used to refer to different information technologies used to collect or process knowledge from humans to produce a considerable information product [1,2]. In recent years, studies have documented that crowdsourcing can produce collective intelligence—i.e., a type of intelligence that emerges through collaboration—and that this intelligence through group collaboration can yield better outcomes [3–7]. To date, several crowdsourcing methods have emerged that allow crowds to collaborate online and produce superior information products. For instance, Wikipedia—a resource based on a crowdsourcing model where many people can contribute and edit encyclopedic articles [3,8]—has been argued to be as accurate as expert-written encyclopedias [9]. Interestingly, there is also evidence showing that, with more editing rounds, the accuracy of Wikipedia improves [10] and results in less biased articles [11].

In the present study, we focus on crowdsourcing in the domain of design—more specifically, that of architectural design. Overall, designing buildings and cities is one of the oldest human activities [12]. With the growth of cities and technological advancements, construction became more complex and involved more artists. This led to a highly collaborative environment in which the architect was in charge of the collaboration. However, during the Renaissance, a radically different approach emerged, which separated the architect from the construction site. As a result, architecture became a fine art, and architects focused on the design of buildings and produced design documents instead of overseeing the construction [13]. Architecture became similar to composing music. The composer writes music for an orchestra, and the architect produces design documents to be constructed by artists and builders. In this way, the profession and nature of collaboration changed dramatically and architecture evolved from a craft to a knowledge industry [13].

Similarly, architecture design competitions are a long-standing tradition [14]. Public architecture competitions were even a part of the Olympic Games until 1948. Such competitions start with an open call—including a design brief—inviting architects to propose designs. After the designs are submitted, a jury or the client select the winning design

from the proposals. Today, architecture is a knowledge industry [15], and computing has revolutionised architectural design [16]. With the emergence and widespread adoption of the Internet, online contests have become popular modalities of architectural competitions and a form of crowdsourcing [17] that can produce collective creativity [18].

It can be argued that, in some ways, design competitions are similar to the creation of Wikipedia. Both are based on a similar principle of making an open call to a large and undefined crowd, collecting information from the crowd, and processing it. However, an important difference between crowdsourcing approaches used to create Wikipedia and those used in design contests is that, while Wikipedia's articles are a product of a collaboration of many participants, in design contests, a collaboration between contestants is not encouraged, so the winning designs are mostly a product of one author (or team). Most of the architectural crowdsourcing websites available today are based on the traditional non-collaborative competition model [17,19]. Given that collective intelligence emerges through interaction among people [20] and correlates with the quality of communication in groups [7], the outcomes of design contests may not be a product of collective intelligence. However, more recent creative crowdsourcing systems, such as OpenIDEO, Quirky and Threadless, have been found to produce collective intelligence through a genuine collaboration of crowds [4]. These systems are called network-based systems since the crowdsourcing tasks in them are interdependent and allow for complex communication [18]. For instance, Dortheimer et al. [21] developed an experimental network-based crowdsourcing method that implements a crowdsourcing collaborative design process.

In this context, the main research questions addressed in this study is as follows: Which kinds of collective intelligence can emerge in a contest- and network- based architectural design crowdsourcing systems? In this respect, we hypothesise that, in a collaborative network-based crowdsourcing system, collective intelligence can emerge due to the interaction between the participants. Conversely, we expect that no collective intelligence would emerge in contest-based crowdsourcing systems. To test these two hypotheses, in this study we analyse workflows in two architectural design crowdsourcing systems: (1) a commercial contest-based system and (2) an experimental collaboration network-based system. The workflow structures are analysed using several collective intelligence models [4] and our own analysis. Then, we identify in which sub-processes collective intelligence can emerge. Finally, we offer a new crowdsourcing process that implements the identified collective intelligence-rich sub-processes.

The contributions that the present study makes to previous research are as follows: (1) we offer a new collective intelligence perspective on online architectural design processes; (2) we discuss evidence on a potential collective intelligence through the analysis and evaluation of collective outcomes; (3) we propose a new crowdsourcing design process that facilitates collective intelligence; and, finally, (4) we argue that previous research of design methods can contribute to the formation of novel crowdsourcing methods.

1.1. The Design Process

While crowdsourcing technology is a relatively recent innovation, the underlying workflow of the design process has been actively investigated since the 1960s [22]. For instance, in his seminal book The Sciences of the Artificial, Herbert Simon presented the design process as a search for a satisfactory design solution [23]. Simon defined design as a rational problem-solving paradigm that emphasises the applicability of design thinking to other disciplines.

The basic structure of the design process was introduced at the Design Methods conference held in 1962. Overall, scholars agreed that a systemic design process consists of analysis, synthesis, and evaluation phases [22,24]. The analysis phase includes collecting, classifying, and mapping the relationships between factors, articulating the problem specifications, and reaching an agreement. Next, in the synthesis phase, creative thinking is applied to perform partial solutions considering limitations. Finally, in the evaluation phase, the solution is evaluated by various evaluation methods. While multiple design

process models have been developed since then, most of them include these three essential stages [25]. However, to date, there is no consensual design process framework [26].

Further research identified several different characteristics of the design process [26]. First, the design process was defined as a process aimed at producing something *novel*. Since creativity is essential for the creation of anything new, it is an integral part of the design process. In this respect, the design process resembles the creative process in cognitive psychology [25]. Accordingly, the analysis, synthesis, and evaluation phases are common to most creative process models and to design processes.

The second important characteristic of the design process is that it is *iterative*. Iterations are essential to progress the design, resolve problems, coordinate, and negotiate design solutions, which suggests inter-dependency of the design task [27]. Through this iterative course, new revisions emerge that either improve the design or are discarded. Accordingly, the repetitive nature of the design process provides a fundamental structure for the development of design process models [28].

The third central characteristic of the design process is that, due to the involvement of multiple stakeholders, numerous tasks, and feedback loops, the design process is also *complex*. With the advancement of the design process through iterations, a network of information is established [29]. This network includes data flow among different collaborators and corresponding feedback loops.

Furthermore, design problems are, in essence, *ill-defined* problems [30]. At the initial stage, the requirements of a problem are not all known. This makes it challenging to synthesise and evaluate corresponding design solutions. Yet, with the progression of the design process, more knowledge is produced and the problem definition becomes more precise. This gives rise to the fourth important characteristic of the design process, namely, it is exploratory, where the solution space co-evolves with the problem space [31]. In applied studies, the co-evolution model was found to be useful in describing the design process [32–34].

Finally, as suggested in previous research on the design process, this process can be engineered to be optimised, and coordinated [35]. Crowdsourcing technologies are based on this explicit workflow process, controlling both the input and specific outputs of each activity. Moreover, the explicit nature of a crowdsourcing process allows one to document communication among activities, measuring its performance and, consequently, optimising the production process [36].

1.2. Crowdsourcing Production Methods

Previous studies have proposed many definitions and classifications of crowdsourcing models [1,37–41]. These classifications were based on different parameters, such as workflow, task kind, crowd selection, incentives, validation, and outcomes of crowdsourcing methods. In this section, we review crowdsourcing workflows, task kinds, crowd selection, incentives, and crowdsourcing examples in the field of architecture and urban design.

1.2.1. Crowdsourcing Workflows

Yu et al. [18] differentiated the following three main kinds of creative crowdsourcing systems: games, contests, and networks. Crowdsourcing *games* can solve computational problems by engaging humans using game mechanisms, such as quizzes, and aggregating the best solutions. Furthermore, in crowdsourcing *contests*, participants generate solutions in parallel in response to an 'open call'. Finally, in *network* crowdsourcing, the problem is divided into pieces that are solved by individual agents and then merged into a solution. All aforementioned three kinds of crowdsourcing systems generate multiple solutions by exploration in parallel. However, games and network approaches are more challenging since they require both a method to break down a task at hand into smaller tasks and a method to further consolidate the pieces into a larger product.

To address this challenge of game- and network-based crowdsourcing systems, several approaches of decomposing tasks and reassembling the results have been developed [37].

First, the *Sequential* workflow divides a task into sub-tasks that are then sequentially solved [42]. In this workflow, each task depends on the output of the previous task. Second, a *parallel* workflow divides the process into sub-tasks that can be performed independently by multiple workers, as in design contests. Third, a *recursive* crowdsourcing workflow is based on the idea that a task is complex and can be subdivided recursively until simple sub-tasks are formulated. Such sub-tasks can be easily performed by doing micro-tasks and then aggregating the outcomes [43]. Fourth, an *iterative* crowdsourcing workflow is based on the concept that complex work is improved by recurring micro-tasks [44]. The input of each micro-task is the previously created work until the work is done or the budget is exhausted. Fifth, a *hybrid* workflow is a combination of several workflow methods that can benefit from some advantages of careful optimisation [45]. Finally, a macro-task workflow is a crowdsourcing process for non-decomposable tasks that require expert knowledge and are largely interdependent [46].

1.2.2. Crowdsourcing Tasks

Crowdsourcing tasks are typically categorised into the following four groups: micro-, complex, macro-, and creative tasks [37]. *Micro-tasks*, which result from the disassembly of a large task, are typically straightforward, independent, time-efficient, and require mere reassembly and aggregation. Typical examples of micro-tasks include image classification or sentence translation tasks. Furthermore, complex tasks, such as writing a paragraph, programming a software function, or proofreading a text, require specific domain knowledge or skills. *Macro-tasks*, which are the third category of tasks, are large expert-level assignments like programming or article writing [46]. Finally, *creative tasks* are contest tasks that require expert-level skills to provide novel ideas and design new solutions. In essence, creative tasks are fundamentally not collaborative. In the field of design, creative tasks are the foundation of design contests.

1.2.3. Crowd Selection

Along with crowdsourcing workflows and tasks, other critical aspects of crowdsourcing are crowd selection and the required expertise level of contributors. In general, contributors in crowdsourcing come from the following two main groups: laypeople and experts [47]. Laypeople are large non-expert crowds that typically contribute by performing simple tasks. In contrast, experts are individuals who possess unique domain knowledge and experience necessary to solve complex problems [48]. For instance, in an architectural project, experts are architects who can design and communicate their solutions. On the other hand, stakeholders in an architectural project are laypeople who cannot be expected to produce design artefacts. However, the project stakeholders possess significant in-depth knowledge of place, environment, and culture, all of which are essential for the evaluation of design solutions [21]. Therefore, their involvement in the design process is essential.

1.2.4. Incentives

Incentives are critical in crowdsourcing and have an enormous impact on output quality. Without incentives, people will mostly not perform work. Incentives are typically categorised into extrinsic and intrinsic. While extrinsic incentives are benefits that result from performing the work, such as money, ratings, or vested interest in outcomes of a project, intrinsic motivations include enjoyment or interest that results from the task itself. There is evidence showing that, while extrinsic incentives are sometimes easier to provide, intrinsic incentives and motivation are critical to achieving good outcomes [49,50].

1.2.5. Crowdsourcing of Architecture and Urban Design

To date, various methods of applying crowdsourcing methods to architecture and design have been proposed. For instance, crowdsourcing methods were applied in participatory design to extract features for design and personal preferences [51,52]. Using parametric design technologies, it was possible to create interactive design tools to allow

crowds to explore design possibilities in urban design [53] and architectural design [54]. Digital sketching software was also used to produce design ideas using various work-flows [55–57]. However, professional architectural design has only been produced using online design contests [17,58,59].

1.3. Collective Intelligence

The first definition of the term collective intelligence was proposed by Pierre Levy [60]. According to Levy's definition, collective intelligence—specifically, in the field of information technologies—is "a form of universally distributed intelligence, constantly enhanced, coordinated in real-time, and resulting in the effective mobilisation of skills" (p. 13). Previous research on collective intelligence is diverse and spans together many disciplines, including psychology [7,20,61], complexity and self-organisation [62], computer science [63], social sciences [64], arts [65], and crowdsourcing [3,4,6]. In all these areas, there are different definitions of collective intelligence that are specific to each of the domains. Accordingly, there is no general and consensual definition of collective intelligence. In this section, we review recent research in collective intelligence and the concept of the wisdom of the crowd.

1.3.1. Wisdom of the Crowd

While the very term collective intelligence is a relatively new coinage, the similar concept of the 'wisdom of crowds' has existed since Aristotle's idea that "many heads are better than one", expressed in his work Politics [66]. In the beginning of the 20th century, Francis Galton analysed farmers' estimations in a weight-judging competition to find that the average weight estimation was more precise than the chance of error of each individual expert [67]. Accordingly, Galton concluded that if an appropriate aggregation method is applied under appropriate conditions, the 'wisdom of crowds' may provide better results than those afforded by experts.

It should be noted, however, that despite important similarities, the concepts of 'wisdom of the crowds' and 'collective intelligence' are not identical. Specifically, 'wisdom of the crowds' is based on collecting and aggregating information from group members to produce better results, and collective intelligence is a complex phenomenon that can give rise to superior intelligence through interaction among group members [4].

1.3.2. Emergence of Collective Intelligence

The conditions that can lead to the formation of a greater degree of intelligence in groups were previously identified in several psychological studies [20]. Specifically, it was found that the most important factors for the emergence of collective intelligence are the quality of communication and diversity of group members [7]. In particular, previous studies found a strong positive relationship between the group's average social sensitivity and collective intelligence. Conversely, when some participants were dominant in group discussions, collective intelligence was lower [20], resulting in group-think [68]. This evidence suggests that central factors that predetermine the emergence of collective intelligence are the availability of the communication network among group members, members' diversity [68], and their social sensitivity [7].

Since collective intelligence is related to communication, the communication network has an important impact on collective intelligence. There is also evidence that good communication is essential for group members to see, copy, and improve ideas [7,68]. However, such communication may have unintended consequences because ideas tend to converge in the early stages, thus reducing the group's diversity of viewpoints and collective intelligence performance [68,69]. It has been suggested that intermittent breaks in communication improve group collective intelligence [70].

On creative crowdsourcing websites, two significant modalities of collective intelligence and wisdom of the crowd were identified [4]. The first one is through 'discussion' that consists of multiple textual expressions on the micro-level and then develops into a discussion on the macro-level. Such a discussion may result in the emergence of consensus.

The second modality is through votes and ratings on the micro-level that later, through aggregation, form a crowd opinion on the macro-level.

Van Du Nguyen and Ngoc Thanh Nguyen suggested the following three measures of collective intelligence in crowdsourcing in estimation tasks [68]: (1) the distance from the collective estimation to the proper value; (2) the number of times when the collective prediction is better than individual predictions; and (3) the quotient between collective error and individual errors. Several experimental studies tested group problem-solving activities compared different group variables to identify which of these factors affect performance [7,20,71]. Furthermore, design collaboration studies measured performance using advertisement click rates [72], novelty measure [73], or subjective evaluations [74,75]. In addition, architectural design has been used in collective intelligence studies as a carefully designed optimisation problem by using a set of building blocks [20,76]. However, evaluating realistic architectural designs is a much more complex, situated, and subjective task. We will address this challenge in the next section.

2. Materials and Methods

In this study, we seek to identify a mechanism that can produce collective intelligence and wisdom of the crowds in architectural design crowdsourcing systems. To this end, we compared the performance of Arcbazar, a commercial architectural contest-based crowdsourcing website [58], and Architasker, an experimental architectural network-based crowdsourcing software constructed and described by Dortheimer et al. [21]. The comparison is based on the analysis of experimental data we previously published elsewhere [77]. These data are useful to learn how the system workflows in the two systems are designed and performed. In the remainder of this section, we first provide a short description of the two systems and then explain how they were studied.

2.1. Contest-Based System

Arcbazar is a commercial architectural competition crowdsourcing website. The website provides a platform for online design contests that are mostly based on a single cycle. Design artefacts are submitted by designers and are later rated by other designers. The 1st, 2nd, and 3rd prize winners are awarded a monetary prize, as determined by the client. However, the client has complete control over the final decision and rating.

To learn how Arcbazar works, we analysed data from a preliminary experiment reported elsewhere [77]. The contest data, including the entries, design brief, questions and answers, and analytical data, are publicly available on the website (see https://www.arcbazar.com/urbanism-plaza-design/competition/design-a-major-public-square-in-jerusalem-israel, accessed on 23 December 2021.) The design contest was launched on 29 June 2018, to document the design process and the crowdsourcing model. The contest brief was taken from another architectural design competition brief for designing the 'Safra Square' in Jerusalem (see https://www.isra-arch.org.il, accessed on 23 December 2021.). The design brief on Arcbazar included the following information: (1) contest terms (payment, winning, and credit); (2) contest goals; (3) project objectives; (4) possible intervention points; (5) tips from the client (i.e., a list of suggestions provided in the original contest brief); (6) a list of notable buildings in Jerusalem; (7) current situation; (8) historical situation; (9) physical description; (10) 28 selected images of the compound; (11) 10 different maps of the area from the municipal website; (12) a 3D area-model; and, finally, (13) the requested architectural artefacts. The total project budget was 1150 USD distributed as follows: 600 USD for 1st place, 300 USD for 2nd place, and 100 USD for 3rd place, while the website fee was 150 USD. The contest produced four designs.

During the time when the project was active on Arcbazar, 25 designers signed up, 13 more saved the project, but only four proposals were submitted. From 1–31 July 2018, the designers worked on their design and could post questions on a public wall (Q&A phase). Seven designers asked eight questions on the competition wall (see Figure 1). Two questions were related to the design requirements, and six further questions were related to technical

issues such as submission deadline, 3D model file format, and the required artefacts. The client published 14 messages answering questions and providing more information. Since the wall is public, the website has a "wall etiquette" message discouraging the designers from discussing individual designs, reviews, or non-related topics.
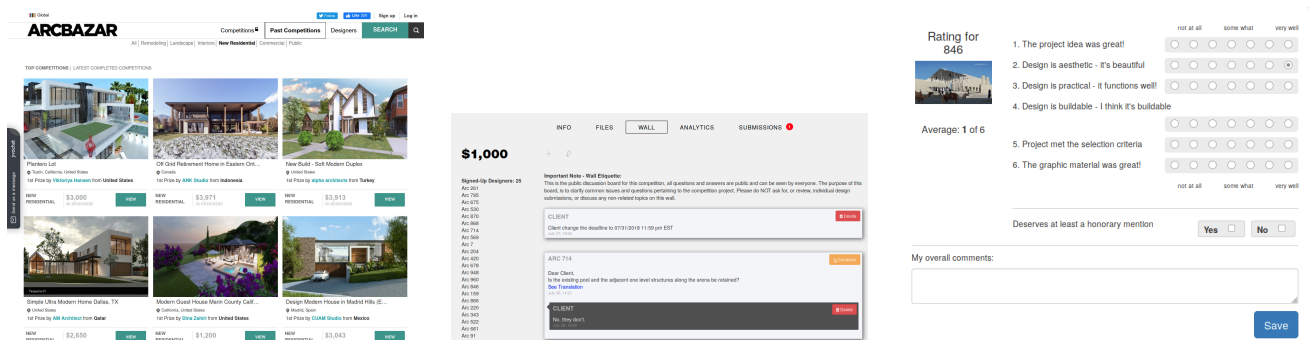

Figure 1. Arcbazar's screenshots: Arcbazar's homepage (**left**). Project discussion wall (**centre**). Design rating interface (**right**).

As mentioned above, after the submission deadline, four designs were submitted. These submissions included multiple images with 3D renders and textual descriptions. This was followed by a public voting session with the participation of several designers, who collectively produced 175 votes. The designs were rated with several statements on a scale from 1 to 7 (see Figure 1). After the voting session, Design 2 received the highest average rating from 13 voters, while the other three designs received lower ratings from fewer voters. Finally, the client chose Design 1 for the 1st place, Design 2 for the second place, and Design 3 for the third place.

2.2. Network-Based System

Architasker is an experimental crowdsourcing software developed by Dortheimer et al. [21] (see Figure 2). The software was developed over two student design workshops at Tel Aviv University. The first workshop aimed to develop the software, while the second workshop's goal was to evaluate the performance of the developed software. In the present study, we used the experimental data from the 2nd workshop, which included one design project, 10 experiments, and 81 design artefacts.

A total of 17 people participated in the design process. Of these, nine were second- to fifth-year architecture students at Tel Aviv University who performed as crowdsourcing workers, and eight participants were professional architects. Among the students, three were in their second year, four were in their third year, one was in their fourth year, and one was in their fifth year. To minimise the potential effect of the workshop's academic requirements on the results, the students were graded based on their attendance and the number of completed micro-tasks. Professional architects were recruited from the 'Upwork' freelance website to participate in several tasks for varying financial compensation.

Architasker's design process started with a design brief that included background information, design requirements, and a 3D CAD model of the area. Several tasks were then assigned to the workforce (i.e., designers, clients, and project stakeholders). The tasks were organised in the following three task sets: design tasks, selection tasks, and review tasks. Each task set included redundant tasks performed in parallel, all of them to be adjourned before moving forward to the next task set.

The first task set included design macro-tasks. For these tasks, designers were provided with the project brief and the 3D model of the planning area. A typical task provided output examples to help obtain necessary results and all steps that had to be followed to complete the task. In advanced design iterations, designers were also provided with previously created artefacts to improve and transform. Furthermore, reviews were presented, along with the artefacts for the designers to relate to. The output of design tasks varied and

could include sketches, as well as 2D and 3D CAD files. Overall, the output of design task sets iterated among hand-drafted sketches, 3D models, and 2D plans.
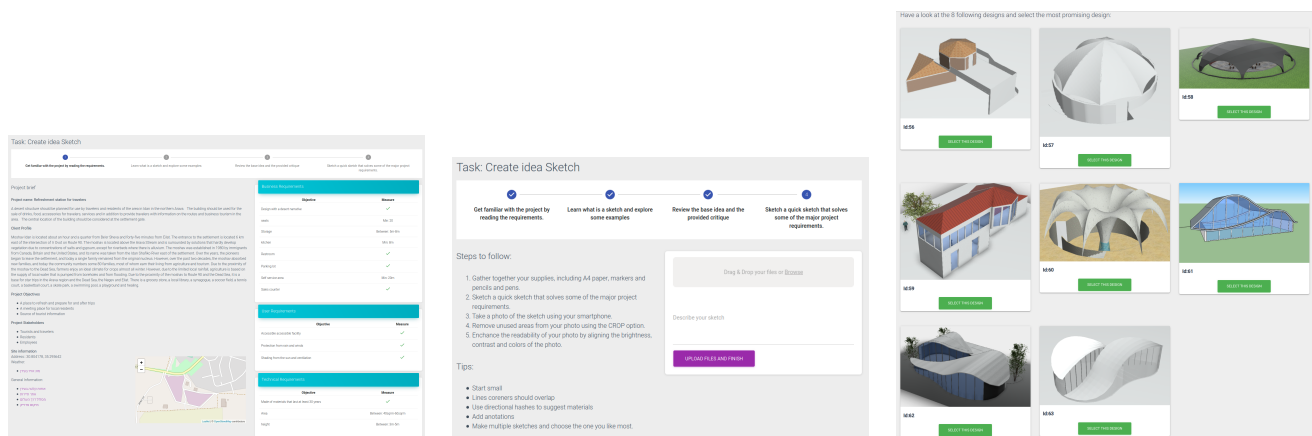

Figure 2. Architasker's screenshots: Design brief screen, with links to local websites in Hebrew (weather, tourism, map, and community) (**left**). Task instructions and upload form (**centre**). Artefact voting screen (**right**).

The second task set consisted of selection micro-tasks. Participants who performed these tasks could be clients, designers, or other project stakeholders. After inspecting the artefacts, the participants were asked to select one artefact that they thought to be the best solution. After all selection tasks in the set were completed, the unselected artefacts were removed, along with at least 50% of the lowest-rated artefacts. The output of the task set was usually one to three artefacts that allowed for further development of different design solutions. The participants provided 54 votes in 10 iterations.

The third task set included review micro-tasks during which textual feedback to the previously selected artefacts was generated from clients, project stakeholders, and designers. The reviews were generated by presenting an architectural artefact to the participants and asking them to answer the following question: "How would you improve this design?" The collected responses were added to the specific artefact object and provided again for a subsequent improvement design task. These tasks were also completed by the participating workshop students, who generated 1642 reviews, with an average of 174.33 reviews per design iteration and 19.13 reviews per artefact over the course of the workshop.

The design brief was as follows: "A new desert tourism centre needs to be planned. It will be used by visitors and residents of the area. The building will be located near the village 'Idan', in the northern Arava desert in Israel. The building should have a store that will sell drinks, food, various products for travellers and provide travellers with information on the routes and businesses in the area. The building will be located at the village gate." The project objectives were to create (1) a place to refresh before and after trips; (2) a meeting place for the local community, and (3) a source of tourist information. The design brief additionally included a list of project stakeholders, site location, an interactive map, business requirements, user requirements, and technical requirements.

2.3. Data Analysis

The research framework we used to compare Arcbazar and Architasker was similar to the one previously proposed by Salminen [4] (see Figure 3). Based on Salminen's conclusions, in order to identify and compare the structure and the emergence of collective intelligence in the analysed systems, we used two analysis tools: the collective intelligence genome and the complex systems approach. We added collective design development analysis, where a participation index is calculated and measures the distance between collaborative and individual performance. Finally, based on the results of the analysis, we

suggested an improved crowdsourcing design process that could arguably facilitate a more collective intelligent process.
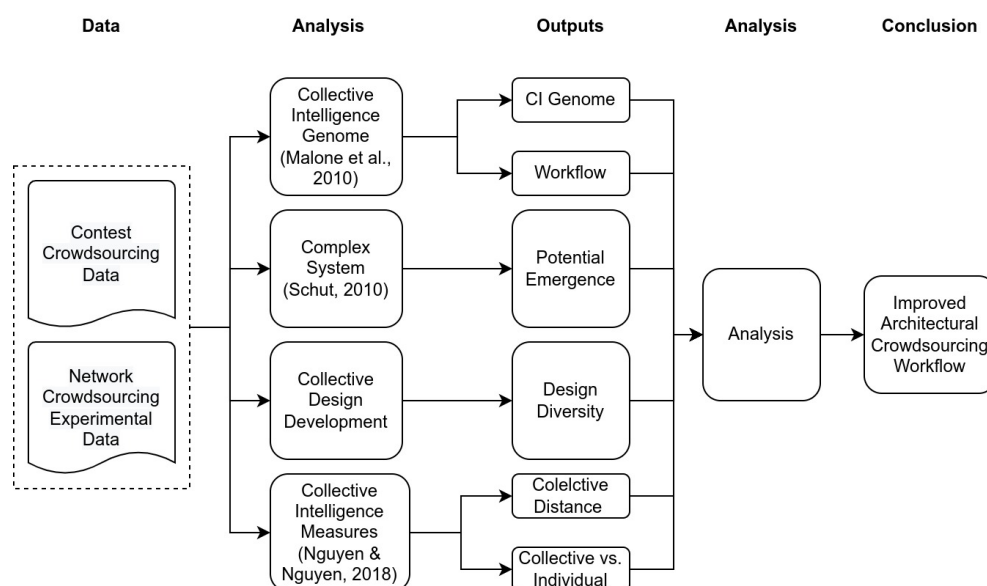

Figure 3. Research process diagram [6,62,68].

2.3.1. Collective Intelligence Genome

The collective intelligence genome [6] is a simple classification system to differentiate between different collective intelligence systems. This system helps one to understand and compare the systems' processes through, first, identification of different phases of production and, second, answering four design questions.

The first question is the goal and what is being done. The possible answers are 'Create' or 'Decide'. Create means that something new, such as a text or a design, was generated. 'Decide' means that the phase aimed to select something, such as a contest winner. The second question is who the staff performing the activity are. This question also has two possible answers: 'Crowd' or 'Hierarchy'. 'Crowd' here refers to a group of undefined participants, while 'Hierarchy' denotes organisers of the process. The third question concerns the incentives and the reasons why actors get engaged in the design process. This question has three possible answers: money, love, or glory. The fourth and final question is related to the structure of the production method—that is, how is it done? The possible answers are collection, contest, collaboration, voting, averaging, and consensus.

Based on the collective intelligence genome, we created workflow diagrams that explain what input was provided and what information was generated at each step. These diagrams helped to identify the communication network that may produce collective intelligence.

2.3.2. Collective Intelligence Complex System

The second analysis method that we used to identify a possible emergence of collective intelligence was based on a complex system approach [62]. Such systems are adaptable and capable of self-organisation, and collective intelligence can emerge under certain conditions.

For the present analysis, we used Schut's [62] complexity-based model to identify systems with potential emergence of collective intelligence. Schut's model is based on the notion that collective intelligence is an emergent phenomenon on a system level that results from an interaction among agents. Shut's model includes the identification of the following two sets of properties: (1) enabling properties and (2) defining characteristics.

Enabling properties, which include adaptivity, interaction, and system rules, are essential properties for the emergence of collective intelligence systems. Adaptivity means that the system is capable of adjusting its structure to a changing environment.

Interaction is thought to occur when there is communication between agents in the system, which enables adequate response to different behaviours. System rules are logical conditions that restrict and adjust information, e.g., task instructions. Since humans are complex agents in crowdsourcing systems based on communication and have explicit rules, enabling properties can be observed on many crowdsourcing websites [4].

Furthermore, defining characteristics are properties that can be recognised in complex systems with collective intelligence. Among others, defining characteristics include local (user) aggregation, global (system) aggregation, randomness, emergence, redundancy, and robustness.

Local aggregation occurs on the individual level, for instance, when a crowd worker composes something creative, such as a review or design. Global aggregation is the ability of the system to adapt itself in response to its environment. In crowdsourcing systems, global aggregation is the sum of votes or a collection of reviews. Furthermore, randomness is a typical element of complex systems identified when there is some random behaviour. For example, in crowdsourcing systems, rated items can be displayed in random order.

Next, emergence refers to the process of local-level aggregations that result in a global level of adaptivity. Whenever emergence occurs, the whole is larger than the sum of its parts. Emergence is the most challenging property in crowdsourcing systems since crowdsourcing involves humans with different behaviour each time, which makes it difficult to predict emergence. Furthermore, redundancy refers to instances when the same information exists or emerges in several places, as when several workers perform a task in parallel. Finally, robustness is related to redundancy and means that, even though some parts of the process can fail, the system will continue to function. For instance, a crowdsourcing system needs to be able to cope with cheaters who perform tasks to get a reward and produce inadequate data.

While most of the parameters briefly reviewed above are observed on many crowdsourcing websites, a significant parameter to identify the emergence of collective intelligence is the local-global aggregation parameter [4]. Both crowdsourcing design systems we analysed are adaptive since they can receive different design challenges and act on them accordingly. They are also interactive and facilitate communication among agents. Finally, there are different rules and constraints, such as restricting votes or participation, in both systems.

In addition, due to human participation, with different individuals involved in each specific case, both systems are characterised by randomness. Similarly, both systems are redundant since all micro-tasks are performed multiple times by multiple people. Finally, the redundancy of the two systems makes them robust so that failures on individual tasks do not lead to system failure. Accordingly, our subsequent analysis focused on the local, global, and emergence parameters.

2.3.3. Collective Design Diversity

Next, we analysed the diversity of design outcomes and produced design development diagrams. Along with documenting the artefacts produced in each design iteration, these diagrams specified which artefacts were selected for further development, showed the dynamics of the design development, and identified how many designers were involved in creating the outcomes.

In order to quantify the level of participation in the resulting design artefact, i.e., the number of unique contributing agents whose designs became part of the final product, the participation index was used. This index allowed us to better understand the level of collaboration, compare collaborative processes, and trace the diversity of contributions over time. Overall, the participation index is a measure of contribution diversity that suggests the potential for the emergence of collective intelligence.

In the present study, the level of collaboration was defined with simple participation index D for an iteration i (see Equation (1)).

$$D_i = n_c - 1 \tag{1}$$

where n_c is the number of contributors (unique participants whose product is actually in the design). Of note, n_c depends on the n, the total number of participants, and i, the number of iterations , since the contributor number cannot be higher than the number of participants or the number of iterations .

$$n_c \leq min([i, n])$$

In the network-based crowdsourcing system, we assumed that the designers would have different skill levels within our heterogeneous group of students and professional architects. Therefore, we hypothesised that some individuals among our study participants would possess better design skills and knowledge. Accordingly, we expected that these individuals' artefacts would be selected more frequently during the design process and that the results would contain the contributions of the best few designers.

2.3.4. Collective Design Measurement

Measuring the emergence of collective intelligence requires a quantitative assessment of the quality of design artefacts produced by various participants during the design process. Since, as discussed previously, architectural designs evaluations are subjective, in the present study, we asked three experts to provide a quantitative evaluation for each artefact. These three experts were professional architects with advanced (Master or Ph.D.) degrees in architecture and experience in educating architecture students. The experts were asked to carefully inspect and then rate the design quality of each artefact on a scale between 1 and 10. The three architects did not know the participants and were affiliated with different universities. The experts' ratings were then normalised, and an average expert score was computed for each artefact.

Following Nguyen and Nguyen [68], based on the experts' evaluations, we computed the following two collective intelligence measures: (1) the distance of the collective prediction from the proper value and (2) the number of times that the collective prediction was better than individual predictions.

For the first measure, we computed the distance of the collective product from the maximal experts' score. This was done by calculating the distance of the average quality of the selected artefacts from the highest-rated artefact in each iteration. Since there were several selected artefacts in each iteration, we computed their average distance from the highest-rated artefacts for each iteration. The smaller the distance was, the better the collective performance was.

Given a collective $X = \{x_1, x_2, \ldots, x_n\}$, which represents individual artefacts, r are the best artefacts and x^* is the collective prediction; in the present study, the distance between the maximal experts' score and the collective produce was defined (see Equation (2)).

$$Diff(X) = 1 - d(r, x^*) \tag{2}$$

The second measure was based on the number of times when the collective performance was better than the performance of individual designers. We calculated the same average distance from the highest-rated artefacts for each participating designer. This was followed by comparing the average collective distance with the average distance of each designer. A smaller distance was assumed to indicate improved performance. If an individual distance was smaller than the collective distance, we interpreted it to mean that an individual performed better than the group. Finally, we also counted the number of designers whose individual performance was higher than the collective value.

Finally, based on the analysis of both crowdsourcing systems, we proposed an improved design process that can facilitate the emergence of collective intelligence in the design process.

3. Results

In this section, we report the results of our analysis of both design crowdsourcing systems.

3.1. Collective Intelligence Genome

In the first analysis, we identified the phases and structure using the collective intelligence genome framework [6]. The contest-based system's genome is shown in Table 1. As can be seen in Table 1, the contest-based system's structure consisted of five phases, where the client (hierarchy) created the challenge (1) and decided on the winner (5). After creating the challenge, the three subsequent phases involved the participation of a crowd of designers.

Table 1. Collective intelligence genome of the contest-based system.

Phase	What		Who	Why	How
1. Challenge	Create	Brief	Hierarchy	Extrinsic	Hierarchy
2. Q&A	Create	Challenge clarification	Crowd and Hierarchy	Extrinsic	Collection
3. Design	Create	Designs	Crowd	Extrinsic	Collection
4. Rating	Create	Average scores	Crowd	Intrinsic	Averaging
5. Winner selection	Decide	Improved artefacts	Hierarchy	Extrinsic	Hierarchy

First, in the Question and Answers (Q&A) phase, the designers asked the client clarification questions about the requirements (2); accordingly, the client's responses led to the creation of a collection of answers. In the next step, complete designs, including numerous architectural artefacts, were anonymously and privately submitted to the client (3). The outcome of this phase was a collection of designs. The submission process was followed by a week-long rating phase when all submitted designs were made public, and all non-participant website designers were able to rate the designs (4). In addition, the client could invite friends and family to participate in the rating phase. This phase produced an average design score that was visible to everyone. Finally, the client received crowd ratings and selected the winners (5). During the winner selection process, the client was not obliged to take into account crowd ratings.

The results of our collective intelligence genome analysis of the contest-based system provide insight into the design process. Specifically, we identified three different subprocesses that involved the crowd producing new information products. First, the questions and answers collections were produced through the interaction between the crowd and the client. The wall served as a shared memory that helped other designers that had similar questions. Second, a collection of designs was produced in parallel, i.e. without any interaction among the designers. Third, a crowd opinion was formed by averaging the crowd's ratings.

Figure 4 shows the design process workflow of the contest-based system. As can be seen in Figure 4, most of the process was linear and included challenge, design, rating, and selection. Thus, the Q&A phase, which occurred during the design phase, included a form of a feedback loop through the shared memory of the wall, which has the potential for complex behaviour and emergence.
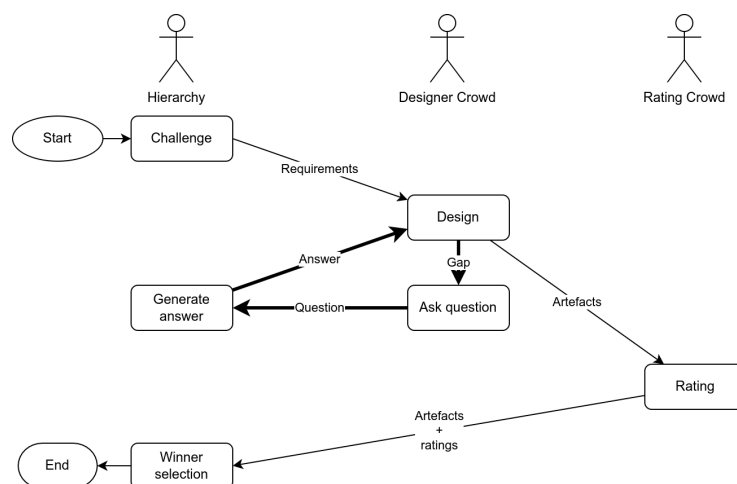
Figure 4. Contest-based system's workflow actor diagram. The design, question, and answer phases appear in bold as they create a feedback loop.

The network-based system's genome starts with a phase similar to the "challenge" phase on the contest-based system (see Table 2). The client provided a brief of the project, which did not involve the crowd (1). In the next phase, the designers, in parallel, produced concept design artefacts that formed a collection of artefacts (2). This process was similar to the contest-based system's process that started with a design brief and a parallel design process. Of note, the design task on the network-based system was considerably shorter and aimed at producing a napkin sketch instead of a complete design solution. The next step was voting: the project stakeholders voted on artefacts and, by aggregating the votes, decided which of these artefacts would be further developed (3). In the subsequent review phase, project stakeholders created a collection of reviews for each selected artefact (4). A stopping condition at this point was when the client (hierarchy) could decide that further improvements were still required and the process could adjourn (5). Upon the continuation of the process, the selected artefacts and reviews became the input of an artefact improvement phase (6). During this phase, the designers created a new and improved collection of artefacts based on the reviews. Finally, the improved artefacts became the input of the selection phase again, forming an iterative process.

Table 2. Collective intelligence genome of the network-based system.

Phase	What		Who	Why	How
1. Challenge	Create	Brief	Hierarchy	Extrinsic	Hierarchy
2. Artefact generation	Create	Artefacts	Crowd	Extrinsic	Collection
3. Selection	Decide	Selection count	Crowd	Extrinsic	Voting
4. Review	Create	How to improve the artefacts	Crowd	Extrinsic	Collection
5. Stopping condition	Decide	Best design	Hierarchy	Extrinsic	Hierarchy
6. Improve artefact	Create	Improved artefacts	Crowd	Extrinsic	Collection

Figure 5 shows the design workflow of the network-based system. While the design process started similarly to the contest-based system, it iterated between design (generation or improvement), selection, and review phases. This iteration used the artefact and reviews as input for the subsequent phase as shared memory, and forced collaboration, and produced a feedback loop. However, unlike in the contest-based system, there was no possibility for the crowd to form a discussion.
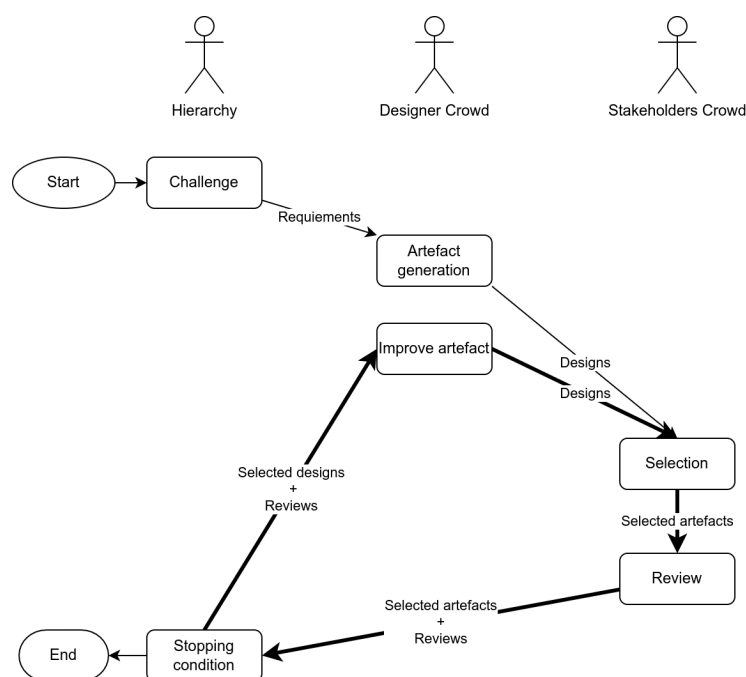
Figure 5. The network-based system's workflow diagram. The design, selection, and review phases rarely appear in bold as they created a feedback loop.

3.2. Collective Intelligence System

In the second analysis, we focused on collective intelligence systems [4,62]. As discussed previously, the local–global aggregations that provide an opportunity to evaluate the emergence of collective intelligence are important indicators in this respect.

The summary of the analysis of the contest-based system is presented in Table 3. The analysis of the contest-based system revealed the potential emergence of collective intelligence through the Q&A (1) sub-process. Individual questions with client answers were aggregated into publicly accessible collections. Such a discussion improved the project's requirements and could have the potential to result in a consensus that previously has not existed. The results also revealed that the average crowd ratings formed a new crowd opinion that also did not previously exist. The aggregated votes were a kind of 'wisdom of the crows', and the Q&A signalled the emergence of collective intelligence.

However, the design sub-process (2) produced a new collection of designs. The sub-process had no mechanism of aggregating or filtering them to produce a consensus. While it produced new information, it missed the emergent property that is essential to form collective intelligence. The artefact development tree of the contest-based system shows that the produced artefacts were the product of a single contest participant (see Figure 6).
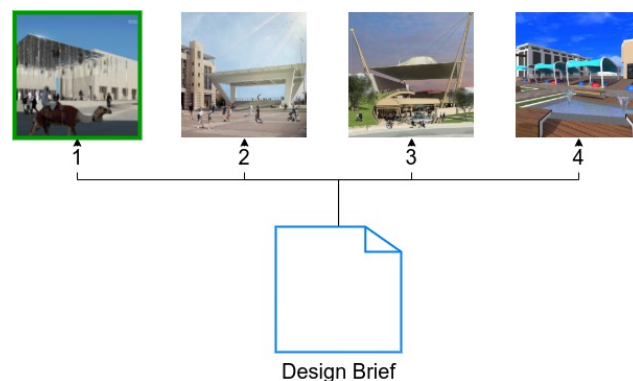

Figure 6. Contest-based system's artefact tree. Each artefact has a number that identifies a specific designer.

Table 3. Local–global defining properties of contest-based system.

Local	Global	Emergence	Kind
1. Q&A	Question and Answers collection	Consensus	Collective Intelligence
2. Design generation	Design collection	No	
3. Individual vote	Aggregated votes	Crowd opinion	Wisdom of the Crowd

Furthermore, the results of our analysis of the network-based system revealed that the network-based artefacts (1) and reviews (2) were created as collections that, in themselves, did not produce an output that could suggest the emergence of collective intelligence or wisdom of the crowd (see Table 4). However, the selection process aggregated individual selections to a crowd opinion that is a form of wisdom of the crowd. Moreover, the iterative improvement sequential micro-task provided the previously created artefacts to the designers for further development. As the process continued, more and more designers contributed their skills and expertise to the collectively produced artefacts.

Table 4. Local–global defining properties of the network-based system.

Local	Global	Emergence	Kind
1. Artefact generation	Artefact collection	No	
2. Review generation	Review collection	No	
3. Artefact improvement	Collection of sequential artefact improvements	Consensus	Collective Intelligence
4. Individual selection	Aggregated selections	Consensus	Wisdom of the Crowd

3.3. Collective Design Diversity

The design artefact development diagram in Figure 7 helps to identify whether the produced artefact resulted from a collective effort. The number of produced designs and the participation index are summarised in Table 5. Interestingly, the resulting artefact after 10 design iterations was a product of six designers ($D = 5$). The work of Designer 9 was selected three times; that of Designers 10 and 4 was selected twice each; and that of Designers 14, 18, and 8 was selected once each. This means that the work of six out of 18 designers was included in the final results, as compared to one contributor ($D = 0$) in the expected results from the contest-based system.

Table 5. Participation index for each design iteration in the network-based crowdsourcing system.

Design Iteration (i)	1	2	3	4	5	6	7	8	9	10
Number of produced designs in each iteration	9	12	10	8	6	6	6	8	8	8
Participation index (D)	0	1	2	3	3	4	5	5	5	5

A different look at the process, instead of the resulting artefact, revealed that nine designers' artefacts were selected during the design process. Furthermore, the artefacts produced by Designer 9 was selected four times; those of Designers 3 and 4 were selected three times; those of Designers 10 and 14 were selected two times; and, finally, those of Designers 2, 8, 13, and 18 were selected once. This result supports our hypothesis that the artefacts of these designers would be more frequently selected in a group with individuals possessing different design skill levels.

Considering that some of our participants had superior design skills, we anticipated that the final output would consist of the contributions made by these expert designers. Interestingly, however, considering that the possible maximum was 10, the results also showed that the final artefacts were joint products of six ($n_c \leq min([i = 10, n = 17])$).

As can be seen in Table 5, the number of contributors increased when the process started and, after the seventh iteration, it became fixed. In the eighth to the tenth iterations, no new contributors were added, meaning that a group of highly skilled crowd workers emerged from the process, outperforming the remaining participants. Since the design

process is redundant, and it is expensive to pay for work that might be discarded, a group of skilled workers can be identified, while the rest of designers can be removed. However, further research would be needed to understand how reducing the number of designs in later stages of the process would affect the quality and performance of the process.
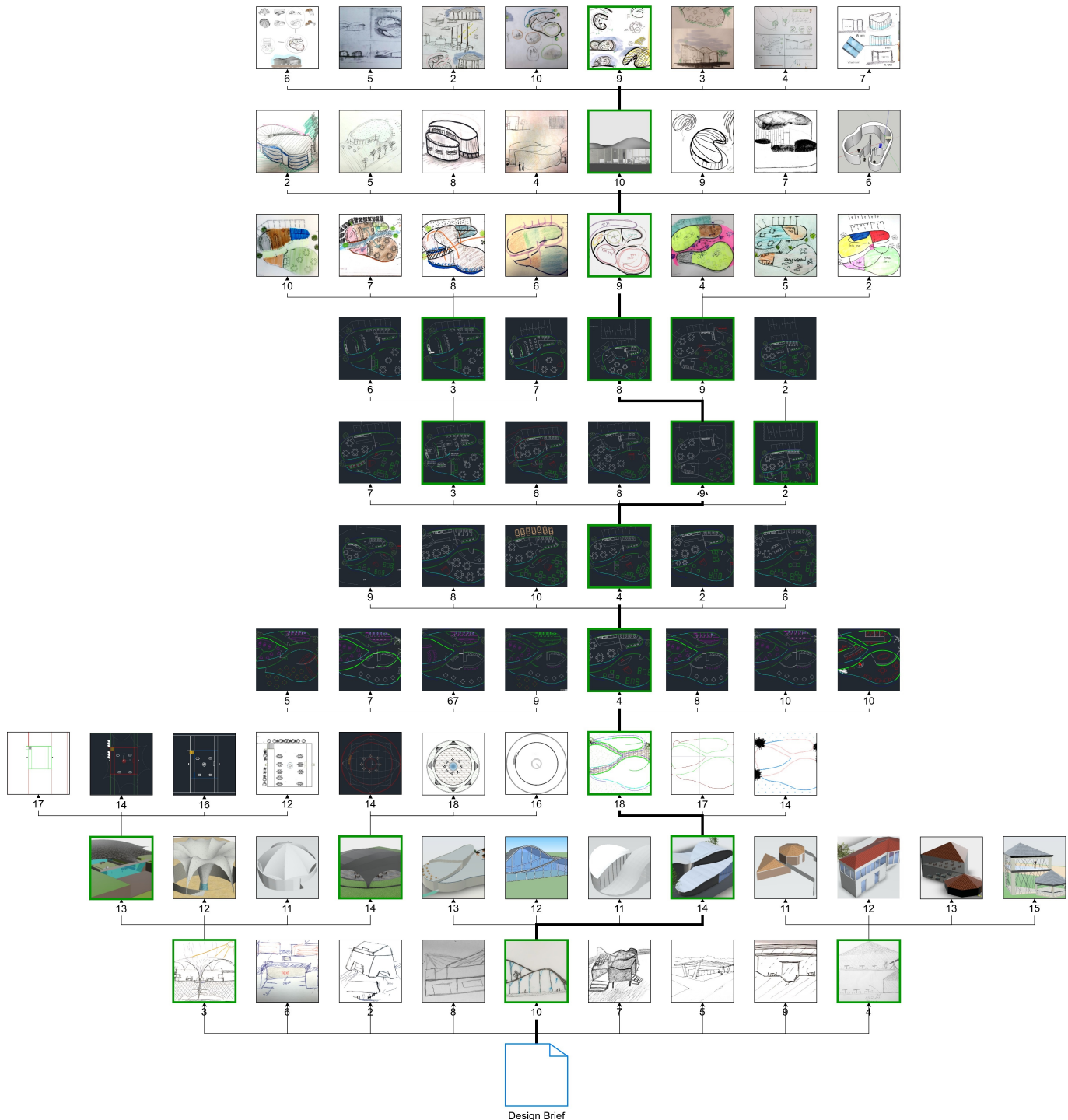

Figure 7. Architasker's artefact tree. Each artefact has a number that identifies a specific designer. The artefacts are ordered in rows representing the corresponding design improvement iterations.

3.4. Collective Design Measurement

Table 6 shows both individual and collective average expert ratings and average distances from the highest-rated artefact, as well as the number of artefacts produced

by each designer. The average distances were calculated following [68] based on expert architects' evaluations. The average distance was the difference between the highest-rated artefact in a specific iteration and the artefact produced by a designer. The collective average distance was the average distance of the artefacts selected by the participants to be further developed in each iteration. The smaller the distance was, the more accurate the artefact was.

Table 6. Collective and individual average distance from the highest-rated artefact in each design iteration.

Iteration	1	2	3	4	5	6	7	8	9	10
Designer 2	1.926	1.926	1.926	2.071	1.117	0.899	0.764	0.799	0.813	0.738
Designer 3	0.000	0.000	0.000	0.144	0.154	0.000	0.138	0.144	0.107	0.366
Designer 4	0.963	0.963	0.963	0.722	0.539	0.578	0.523	0.385	0.608	0.527
Designer 5	0.193	0.193	0.193	0.433	0.443	0.482	0.427	0.722	0.926	0.848
Designer 6	2.697	2.697	2.697	1.686	1.310	1.204	1.024	1.236	1.208	1.064
Designer 7				0.530	0.539	0.578	0.523	0.385	0.608	0.751
Designer 8	1.252	1.252	1.252	1.236	1.069	1.040	0.876	0.888	0.926	0.896
Designer 9	0.963	0.963	0.963	0.915	0.668	0.819	0.754	0.690	0.795	0.751
Designer 10	0.578	0.578	0.578	0.337	0.250	0.289	0.234	0.222	0.203	0.352
Designer 11		0.899	0.899	1.044	1.053	1.092	1.037	1.044	1.006	0.976
Designer 12		0.963	0.867	1.011	1.021	1.060	1.005	1.011	0.974	0.944
Designer 13		1.349	1.349	1.493	1.503	1.541	1.486	1.493	1.456	1.426
Designer 14		0.096	0.433	0.578	0.588	0.626	0.571	0.578	0.540	0.511
Designer 15		1.541	1.541	1.686	1.695	1.734	1.679	1.686	1.648	1.618
Designer 16			1.252	1.397	1.406	1.445	1.390	1.397	1.359	1.329
Designer 17			1.220	1.365	1.374	1.413	1.358	1.365	1.327	1.297
Designer 18			0.289	0.433	0.443	0.482	0.427	0.433	0.396	0.366
Collective distance	0.514	0.321	0.214	0.353	0.321	0.375	0.349	0.353	0.314	0.283
High performing individuals	2	3	2	2	2	2	2	2	2	0

The results revealed that the collective distance was 0.28, suggesting that a group outperformed each individual participant. This was a surprising finding because in the previous analysis the participation index settled on 5, and we expected one of those participants (i.e., Designers 8–10, 14, 18) to demonstrate a stable performance superior to that of the group. Furthermore, the collective distance demonstrated an improvement trend throughout the design process.

Additionally, the results show that until the 9th iteration, there were between two and three designers with better individual performance than the collective performance, specifically Designers 3 and 10. Their higher performance can be explained by the fact that they were both senior students. However, on the 10th iteration, Designers 3 and 10 under performed, and the collective score was overtaken. Moreover, as mentioned before, the artefacts of Designer 3 were selected three times during the design process. However, the resulting artefact did not include even one artefact produced by Designer 3, suggesting that the collective process is not an outcome of the most senior designers but offers a more diverse process to include designers' ideas with various skill levels.

Furthermore, during iteration 6, the crowd selected the design artefact with the lowest expert rating. However, while the artefact was rated low, the crowd identified it to have unique qualities. Once the design was improved on the seventh iteration, the resulting artefact was evaluated as the highest by the experts since it solved architectural programme issues. Consequently, it became a part of the design process outcome.

4. Discussion

Our analysis of the two architectural crowdsourcing systems revealed potential processes that could drive the emergence of collective intelligence and wisdom of the crowd. Specifically, we identified the following three kinds of mechanisms: (1) online design discussions, such as questions and answers, which can produce collective intelligence; (2) sequential design improvements, which can produce a collaborative design; and (3) voting and rating of designs, which can give rise to the wisdom of the crowd. These observations were supported with empirical evidence showing that the crowdsourcing process performed better than individual designers. Based on these results, we proposed an improved collaborative design process.

4.1. Design Discussions

Previous research has documented that design discussions may be useful to produce collective intelligence, explore the problem space, identify design ideas, and evaluate design solutions [4]. For design requirement clarifications, the contest-based system used an online discussion using a "wall" as collective memory. However, since the design process is based on a contest, the participants were discouraged from sharing their design ideas or developing discussions to limit the possibility of copying ideas. This makes sense in a competition, and the competitors were unwilling to provide their competitors with an advantage by sharing their ideas. Consequently, due to competition, collective intelligence did not emerge .

4.2. Sequential and Parallel Design Development

Through the network-based crowdsourcing system analysis, we identified potential collective intelligence emergence through the iterative design process. This was supported by evidence that, after 10 design iterations, the collective distance measure outperformed the individuals' distance measure. The final design, produced by six different designers, was diverse. Based on these findings, it can be concluded that the iterative hybrid workflow (sequential and parallel) can produce a design that is a product of collective intelligence (see Table 6 and Figure 8). Accordingly, the outcomes of the network-based crowdsourcing system are arguably similar to those produced by Wikipedia contributors, where several authors sequentially improve the outcome and conduct discussions.
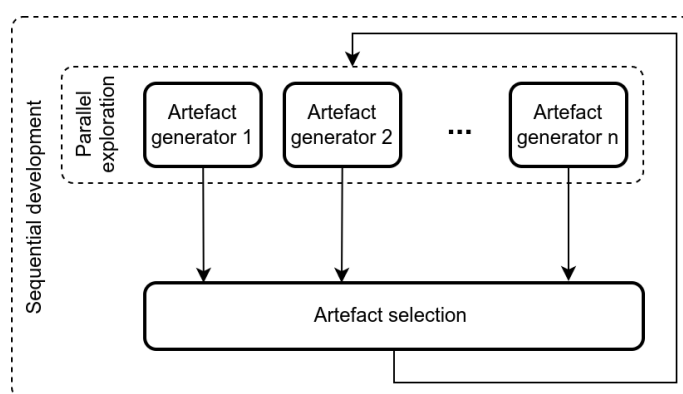

Figure 8. Hybrid design crowdsourcing development workflow. Design is explored in parallel in the artefact generator processes. The artefact generator outcomes are filtered in the selection process. The best artefacts are again provided to a sequential artefact generator process.

The iterative process facilitated a feedback loop, which is an essential quality of collective intelligent systems [4,62]. The feedback loop exposes all produced artefacts to the designers, allowing them to reflect on their designs and merge features from other designs into a new artefact. Therefore, each generation of artefacts is based not only on the best artefacts previously selected by the crowd but also on a possible combination inspired by

other designers. Such production of artefacts based on various kinds of collaboration is not possible in contest-based crowdsourcing systems.

Another essential aspect is the parallel exploration that can improve collective intelligence as an integral part of the hybrid workflow. Previously, breaks in the interactions were reported to improve collective intelligence [70] because, in a well-connected social network, social influence can undermine the wisdom of the crowd [69]. Therefore, parallel design tasks followed by sequential improvement would produce a higher degree of collective intelligence than in the case when parallel tasks are not performed collaboratively.

One of the questions that arises from an analysis of design development dynamics is the convergence of design and the establishment of the design contributor group. On the one hand, a design process aims to produce a monovalent solution, so there must be a convergence of the design into a single solution. On the other hand, there is a risk that an early convergence of design is an expression of group-thinking that reduces the crowd's collective creative ability. Unfortunately, the data from this study are not sufficient to study this phenomenon. A rigorous study is needed that compares different communication modalities, varying participant group sizes, and the convergence of the process, as discussed further in the future research section.

4.3. Evaluation and Selection

Selection and rating processes are challenging in the domain of design—an area where there are no "correct" answers and where one would rather talk about a good 'fit' [23]. In addition, design evaluation may considerably vary even among experts. In this context, it is pivotal that the votes of the crowd make sense and help bring forth the fittest designs.

In our results, we observed that both systems used 'wisdom of the crowd' rating and voting mechanisms. These mechanisms are straightforward to implement and helpful in evaluating and selecting the best designs. However, in both crowdsourcing systems, the limitation of the critical selection tasks was that these tasks are susceptible to abuse, bias, or cheating. According to our findings, both systems addressed these challenges. In the contest-based system, these challenges were mitigated since the rating did not directly affect the process outcome. Instead, the rating was provided to the client and served as a recommendation for the client's selection of winners.

In the network-based system, another approach that could facilitate a higher degree of collective intelligence was used. Specifically, the selection sub-process resulted in a selection of several design artefacts rather than just one winner. By retaining several artefacts, the design process became more diverse, allowing for the parallel exploration of several options. Such diversity is essential for the emergence of higher collective intelligence [20]. Finally, in both systems, a selection of designs was given to the client or project stakeholders, who would then live with the consequences of their decision.

4.4. A New Crowdsourcing Design Workflow

Crowdsourcing workflows can considerably benefit from design methods research. Creative crowdsourcing systems [4,8] consist of a workflow that can be distinctively divided into several sub-processes connected and governed by algorithms regulating the input, output, and execution of those sub-processes. Researchers can experiment with different creative workflow processes using crowdsourcing, measuring, and optimising their performance [36]. This, in turn, can produce computer systems that manage crowds, thus producing design and other outcomes of superior quality.

Based on the results of the present study, we proposed a new collaborative "Design Method", which can be implemented as a crowdsourcing design workflow process. Specifically, we argued that an iterative process consists of the following three iterative sub-processes: (1) discussion, (2) parallel design synthesis, and (3) selection (see Figure 9).

1. In the discussion stage, designers share ideas about design requirements and potential ideas. Since discussions are performed in natural language (rather than sketches), they allow project stakeholders and clients to better articulate design requirements with

the assistance of the participating designers. The output of this stage is a conversation that can be summarised into an improved design brief.

2. In the during *parallel synthesis*, designers produce sketches and diagrams of artefacts providing a solution to the design problem based on the design brief. This exploration should yield a diversity of preliminary design sketches for further discussion and elaboration. To ensure the diversity of the proposed design solutions, the designers should work in parallel, i.e. with limited communication among them. The outcome of this stage would be a collection of design artefacts.

3. In the *selection* stage, design artefacts are subjectively evaluated by the project stakeholders and designers to identify the most promising designs. The most straightforward way to select designs is by voting on the best designs. The aggregated votes would then help to identify the designs that should be removed from the process, leaving a sub-set of the fittest designs to allow diversity.
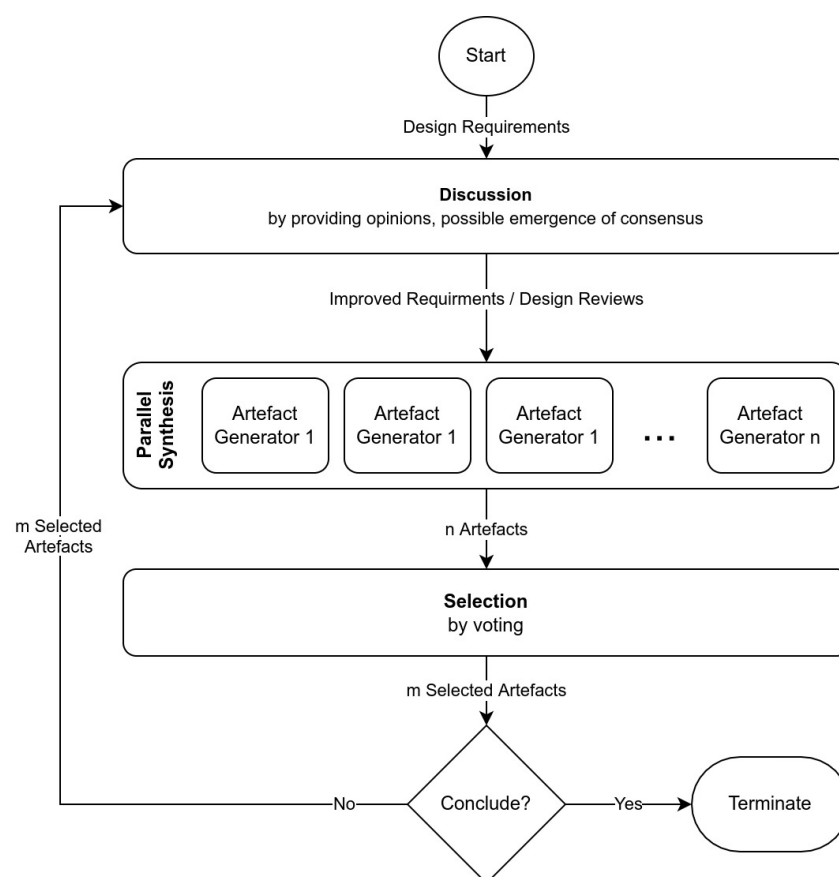

Figure 9. The new crowdsourcing design workflow. The process is made out of discussion, parallel design synthesis, and selection sub-processes.

Once the sub-process adjourns, it restarts with a new discussion where all project stakeholders and designers would provide, in open discussion, both ideas for the improvement of and a critique of the selected design artefacts. Thereupon, the review conversation and artefacts would be provided to the parallel synthesis phase again. The process would adjourn by an external process manager's decision when the process was exhausted.

The components of the proposed collaborative design process are similar to those of the "systematic design process" Jones presented in 1962 at the Design Methods conference [24]. In addition, it can be argued that the proposed process implements Maher et al.'s co-evolution model, which explores the problem space and the solution space iteratively [31]. In the presented process, the problem space is adapted at the discussion sub-process, while the solution space is explored at the parallel synthesis and selection sub-processes.

The aforementioned collaborative process can potentially offer better performance than a contest. First, it involves project stakeholders who navigate the creative process, thereby offering a democratic method of participatory design. Second, it allows many designers to participate without any financial risk involved in participating in a competition. Third, it is scalable and allows many people to participate and improve the design, contribute their intelligence, and make the process more diverse and democratic.

4.5. Limitations

The present study has several limitations. First, expert evaluations performed on this study were subjective, which limits the reproducibility of our findings. Second, since no collective intelligence measuring method for design is available yet, we adjusted collective intelligence measures from previous studies that measured crowds' ability to predict variables.

4.6. Future Research

The present study leaves several open questions that should be addressed in further research. For instance, in the present study, we did not compare the design quality obtained via a network-based crowdsourcing system with the outcomes of a contest-based system. While our results indicate that the collective performed better than the individual participants, it remains to be established whether such a collaborative outcome would be superior to an outcome of a contest. Furthermore, in future research, the new crowdsourcing workflow would need to be measured, and the impact of discussions on the quality of the outcomes of the design process would need to be evaluated. To this end, in our further research, we plan to compare the network and the contest-based crowdsourcing systems by providing them with an identical design brief and evaluating the outcome design quality by specialists.

Additionally, the results showed that the participation index converged after several iterations, which raised research questions. Future research will investigate the relationship between the convergence and the emergence of collective intelligence and identify the various parameters that affect the convergence (such as the number of participants and number of iterations). A study of this nature will also identify the minimum number of participants required to achieve collective intelligence and the minimum number of design iterations when collective intelligence occurs.

5. Conclusions

The present investigation is one of the first studies to explore the emergence of collective intelligence in crowdsourcing for architectural design. Specifically, we compared manifestations of collective intelligence in two architectural design crowdsourcing systems: a commercially available contest-based crowdsourcing system and a network-based collaborative design crowdsourcing system. Previous research on collective intelligence in crowdsourcing systems revealed that collective intelligence manifests itself via different kinds of discussions and voting. In this respect, our findings showed that architectural design crowdsourcing systems could successfully exploit the wisdom of the crowd with voting micro-tasks for design evaluation. Moreover, we also observed that network-based hybrid (parallel and sequential) workflows could produce collective intelligent products. Overall, collaborative network-based crowdsourcing systems are promising tools that can be effectively used to integrate project stakeholders' knowledge into different parts of the design process and thus produce genuinely collaborative designs.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to experimental data used for this study being published elsewhere, and research associates provided the expert evaluations.

References

1. Brabham, D.C. Crowdsourcing as a Model for Problem Solving. *Converg. Int. J. Res. New Media Technol.* **2008**, *14*, 75–90. [CrossRef]
2. Howe, J. The Rise of Crowdsourcing. *Wired Mag.* **2006**, *14*, 1–4.
3. Leimeister, J.M. Collective Intelligence. *Bus. Inf. Syst. Eng.* **2010**, *2*, 245–248. [CrossRef]
4. Salminen, J. The Role of Collective Intelligence in Crowdsourcing Innovations. Ph.D. Thesis, Lappeenranta University of Technology, Lappeenranta, Finland, 2015.
5. Shen, H.; Li, Z.; Liu, J.; Grant, J.E. Knowledge Sharing in the Online Social Network of Yahoo! Answers and Its Implications. *IEEE Trans. Comput.* **2014**, *64*. [CrossRef]
6. Malone, T.W.; Laubacher, R.; Dellarocas, C. The collective intelligence genome. *IEEE Eng. Manag. Rev.* **2010**, *38*, 38–52. [CrossRef]
7. Engel, D.; Woolley, A.W.; Jing, L.X.; Chabris, C.F.; Malone, T.W. Reading the Mind in the Eyes or Reading between the Lines? Theory of Mind Predicts Collective Intelligence Equally Well Online and Face-To-Face. *PLoS ONE* **2014**, *9*, e115212. [CrossRef]
8. Malone, T.W.; Laubacher, R.; Dellarocas, C. Harnessing crowds: Mapping the genome of collective intelligence. *MIT Sloan Sch. Manag.* **2009**, *1*, 1–20. [CrossRef]
9. Jim Giles. Internet encyclopaedias go head to head. *Nature* **2005**, *438*, 900–901. [CrossRef]
10. Wilkinson, D.M.; Huberman, B.A. Cooperation and quality in Wikipedia. In Proceedings of the Conference on Object-Oriented Programming Systems, Languages, and Applications (OOPSLA), Phoenix, AZ, USA, 21–25 October 2007; pp. 157–164. [CrossRef]
11. Greenstein, S.; Zhu, F. Do experts or crowd-based models produce more bias? Evidence from encyclopedia britannica and wikipedia. *MIS Q. Manag. Inf. Syst.* **2018**, *42*, 945–958. [CrossRef]
12. Vitruvius, P.M. *The Architecture of Marcus Vitruvius Pollio*; Lockwood & Co.: London, UK, 1874.
13. Carpo, M. *The Alphabet and the Algorithm*; MIT Press: Cambridge, MA, USA, 2011.
14. Lipstadt, H. Can 'art Professions' Be Bourdieuean Fields Of Cultural Production? The Case Of The Architecture Competition. *Cult. Stud.* **2003**, *17*, 390–419. [CrossRef]
15. Porat, M.U. *The Information Economy: Definition and Measurement*; Technical Report; Office of Telecommunications (DOC): Washington, DC, USA, 1977.
16. Wright Steenson, M. *Architectural Intelligence: How Designers and Architects Created the Digital Landscape*; The MIT Press: Cambridge, MA, USA, 2017; p. 328.
17. Kamstrup, A. Crowdsourcing and the Architectural Competition as Organisational Technologies. Ph.D. Thesis, Copenhagen Business School, Frederiksberg, Denmark, 2017.
18. Yu, L.; Nickerson, J.V.; Sakamoto, Y. Collective Creativity: Where we are and where we might go. In Proceedings of the Collective Intelligence 2012, Cambridge, MA, USA, 18–20 April 2012. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2037908 (accessed on 23 December 2021).
19. Angelico, M.; As, I. *Crowdsourcing Architecture: A Disruptive Model in Architectural Practice*; ACADIA: San Francisco, CA, USA, 2012; pp. 439–443.
20. Woolley, A.W.; Chabris, C.F.; Pentland, A.; Hashmi, N.; Malone, T.W. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* **2010**, *330*, 686–688. [CrossRef] [PubMed]
21. Dortheimer, J.; Neuman, E.; Milo, T. A Novel Crowdsourcing-based Approach for Collaborative Architectural Design. In *Anthropologic: Architecture and Fabrication in the Cognitive Age—Proceedings of the 38th eCAADe Conference*; Education and Research in Computer Aided Architectural Design in Europe: Berlin, Germany, 2020; Volume 2, pp. 155–164.
22. Cross, N. *The Automated Architect*; Pion Limited: London, UK, 1977; p. 178.
23. Simon, H.A. *The Sciences of the Artificial*; MIT Press: Cambridge, MA, USA, 1969; p. 123.
24. Jones, C.J. A method of systemic design. In *Design Methods*; Jones, C.J., Thornley, D.G., Eds.; Pergamon Press: Oxford, UK, 1963; pp. 53–73.
25. Howard, T.; Culley, S.; Dekoninck, E. Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Des. Stud.* **2008**, *29*, 160–180. [CrossRef]
26. Wynn, D.C.; Clarkson, P.J. Process models in design and development. *Res. Eng. Des.* **2018**, *29*, 161–202. [CrossRef]
27. Luckman, J. An Approach to the Management of Design. *J. Oper. Res. Soc.* **1967**, *18*, 345. [CrossRef]
28. Takeda, H.; Veerkamp, P.; Tomiyama, T.; Yoshikawa, H. Modeling design processes. *AI Mag.* **1990**, *11*, 37–48.
29. Maurer, M. *Complexity Management in Engineering Design—A Primer*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 1–153. [CrossRef]
30. Simon, H.A. The structure of ill structured problems. *Artif. Intell.* **1973**, *4*, 181–201. [CrossRef]

31.　Maher, M.L.; Poon, J.; Boulanger, S. Formalising Design Exploration as Co-evolution: A Combined Gene Approach. In *Advances in Formal Design Methods for CAD: Proceedings of the IFIP WG5.2 Workshop on Formal Design Methods for Computer-Aided Design*; Springer, Boston, MA, USA, 1996; pp. 3–30. [CrossRef]

32.　Dorst, K.; Cross, N. Creativity in the design process: Co-evolution of problem–solution. *Des. Stud.* **2001**, *22*, 425–437. [CrossRef]

33.　Maher, M.; Tang, H.H. Co-evolution as a computational and cognitive model of design. *Res. Eng. Des.* **2003**, *14*, 47–64. [CrossRef]

34.　Wiltschnig, S.; Christensen, B.T.; Ball, L.J. Collaborative problem–solution co-evolution in creative design. *Des. Stud.* **2013**, *34*, 515–542. [CrossRef]

35.　Browning, T.R.; Ramasesh, R.V. A Survey of Activity Network-Based Process Models for Managing Product Development Projects. *Prod. Oper. Manag.* **2009**, *16*, 217–240. [CrossRef]

36.　Maher, M.L. Design Creativity Research: From the Individual to the Crowd. In *Design Creativity 2010*; Springer: London, UK, 2011; pp. 41–47. [CrossRef]

37.　Bhatti, S.S.; Gao, X.; Chen, G. General framework, opportunities and challenges for crowdsourcing techniques: A Comprehensive survey. *J. Syst. Softw.* **2020**, *167*, 110611. [CrossRef]

38.　Estellés-Arolas, E.; González-Ladrón-de Guevara, F. Towards an integrated crowdsourcing definition. *J. Inf. Sci.* **2012**, *38*, 189–200. [CrossRef]

39.　Hosseini, M.; Shahri, A.; Phalp, K.; Taylor, J.; Ali, R. Crowdsourcing: A taxonomy and systematic mapping study. *Comput. Sci. Rev.* **2015**, *17*, 43–69. [CrossRef]

40.　LaToza, T.D.; van der Hoek, A. Crowdsourcing in Software Engineering: Models, Motivations, and Challenges. *IEEE Softw.* **2016**, *33*, 74–80. [CrossRef]

41.　Nakatsu, R.T.; Grossman, E.B.; Iacovou, C.L. A taxonomy of crowdsourcing based on task complexity. *J. Inf. Sci.* **2014**, *40*, 823–834. [CrossRef]

42.　Jiang, H.; Matsubara, S. Efficient Task Decomposition in Crowdsourcing. In *PRIMA 2014: Principles and Practice of Multi-Agent Systems*; Dam, H.K., Pitt, J., Xu, Y., Governatori, G., Ito, T., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 65–73.

43.　Kulkarni, A.; Can, M.; Hartmann, B. Turkomatic: Automatic Recursive Task and Workflow Design for Mechanical Turk. In Proceedings of the Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 7–11 August 2011; pp. 2053–2058. [CrossRef]

44.　LaToza, T.D.; Ben Towne, W.; Adriano, C.M.; Van Der Hoek, A. Microtask programming: Building software with a crowd. In Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14), Honolulu, HI, USA, 5–8 October 2014; pp. 43–54. . [CrossRef]

45.　Kittur, A.; Smus, B.; Kraut, R. CrowdForge Crowdsourcing Complex Work. In *Human Factors in Computing Systems*; ACM: Santa Barbara, CA, USA, 2011; pp. 43–52. [CrossRef]

46.　Retelny, D.; Robaszkiewicz, S.; To, A.; Lasecki, W.S.; Patel, J.; Rahmati, N.; Doshi, T.; Valentine, M.; Bernstein, M.S. Expert crowdsourcing with flash teams. In Proceedings of the 27th annual ACM symposium on User Interface Software and Technology, Honolulu, HI, USA, 5–8 October 2014; ACM: New York, NY, USA, 2014; pp. 75–85. [CrossRef]

47.　Kittur, A.; Nickerson, J.V.; Bernstein, M.; Gerber, E.; Shaw, A.; Zimmerman, J.; Lease, M.; Horton, J. The Future of Crowd Work. In Proceedings of the 2013 Conference on Computer Supported Cooperative Work—CSCW '13, San Antonio, TX, USA, 21–25 September 2013; ACM Press: New York, NY, USA, 2013; p. 1301. [CrossRef]

48.　Dortheimer, J.; Margalit, T. Open-source architecture and questions of intellectual property, tacit knowledge, and liability. *J. Archit.* **2020**, *25*, 276–294. [CrossRef]

49.　Cerasoli, C.P.; Nicklin, J.M.; Ford, M.T. Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychol. Bull.* **2014**, *140*, 980–1008. [CrossRef]

50.　Rogstadius, J.; Kostakos, V.; Kittur, A.; Smus, B.; Laredoc, J.; Vukovic, M. An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets. In Proceedings of the International AAAI Conference on Web and Social Media, Barcelona, Spain, 17–21 July 2011; Volume 5, pp. 321–328.

51.　Hosio, S.; Goncalves, J.; Kostakos, V.; Riekki, J. Crowdsourcing Public Opinion Using Urban Pervasive Technologies: Lessons From Real-Life Experiments in Oulu. *Policy Internet* **2015**, *7*, 203–222. [CrossRef]

52.　Lu, H.; Gu, J.; Li, J.; Lu, Y.; Müller, J.; Wei, W.; Schmitt, G. Evaluating urban design ideas from citizens from crowdsourcing and participatory design. In Proceedings of the CAADRIA 2018—23rd International Conference on Computer-Aided Architectural Design Research in Asia: Learning, Prototyping and Adapting, Beijing, China, 17–19 May 2018; Volume 2, pp. 297–306.

53.　Birch, D.; Simondetti, A.; Guo, Y.k. Crowdsourcing with online quantitative design analysis. *Adv. Eng. Inform.* **2018**, *38*, 242–251. [CrossRef]

54.　Fisher-gewirtzman, D.; Polak, N. Integrating Crowdsourcing & Gamification in an Automatic Architectural Synthesis Process. In Proceedings of the 36th eCAADe Conference, Lodz, Poland, 19–21 September 2018; Volume 1, pp. 439–444.

55.　Sun, L.; Xiang, W.; Chen, S.; Yang, Z. Collaborative sketching in crowdsourcing design: A new method for idea generation. *Int. J. Technol. Des. Educ.* **2015**, *25*, 409–427. [CrossRef]

56.　Xiang, W.; Sun, L.Y.; You, W.T.; Yang, C.Y. Crowdsourcing intelligent design. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 126–138. [CrossRef]

57.　Yu, L.; Sakamoto, Y. Feature selection in crowd creativity. *Lect. Notes Comput. Sci.* **2011**, *6780 LNAI*, 383–392. [CrossRef]

58. As, I.; Nagakura, T. Crowdsourcing the Obama presidental center. In Proceedings of the Disciplines and Disruption—Proceedings Catalog of the 37th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA), Cambridge, MA, USA, 2–4 November 2017; pp. 118–127.
59. As, I. Competitions in a Networked Society: Crowdsourcing Collective Deisgn Intelligence. In *BLACK BOX: Articulating Architecture's Core in the Post-Digital Era*; ASCA: Pittsburgh, PA, USA, 2019; pp. 268–273.
60. Levy, P. *Collective Intelligence: Mankind's Emerging World in Cyberspace*; Perseus Books: New York, NY, USA, 1997; p. 255.
61. Engel, D.; Woolley, A.W.; Aggarwal, I.; Chabris, C.F.; Takahashi, M.; Nemoto, K.; Kaiser, C.; Kim, Y.J.; Malone, T.W. Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 3769–3778. [CrossRef]
62. Schut, M.C. On model design for simulation of collective intelligence. *Inf. Sci.* **2010**, *180*, 132–155. [CrossRef]
63. Lévy, P. From social computing to reflexive collective intelligence: The IEML research program. *Inf. Sci.* **2010**, *180*, 71–94. [CrossRef]
64. Mačiulienė, M.; Skaržauskienė, A. Emergence of collective intelligence in online communities. *J. Bus. Res.* **2016**, *69*, 1718–1724. [CrossRef]
65. Casal, D.P. Crowdsourcing the Corpus: Using Collective Intelligence as a Method for Composition. *Leonardo Music J.* **2011**, *21*, 25–28. [CrossRef]
66. Landemore, H. Collective Wisdom. In *Collective Wisdom*; Landemore, H., Elster, J., Eds.; Cambridge University Press: Cambridge, UK, 2012; pp. 1–20. [CrossRef]
67. Galton, F. Vox Populi. *Nature* **1907**, *75*, 450–451. [CrossRef]
68. Nguyen, V.D.; Nguyen, N.T. Intelligent Collectives: Theory, Applications, and Research Challenges. *Cybern. Syst.* **2018**, *49*, 261–279. [CrossRef]
69. Lorenz, J.; Rauhut, H.; Schweitzer, F.; Helbing, D. How social influence can undermine the wisdom of crowd effect. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9020–9025. [CrossRef]
70. Bernstein, E.; Shore, J.; Lazer, D. How intermittent breaks in interaction improve collective intelligence. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 8734–8739. [CrossRef]
71. Kim, Y.J.; Engel, D.; Woolley, A.W.; Lin, J.Y.T.; McArthur, N.; Malone, T.W. What Makes a Strong Team?: Using Collective Intelligence to Predict Team Performance in League of Legends In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, OR, USA, 25 February–1 March 2017; ACM: New York, NY, USA, 2017; pp. 2316–2329. [CrossRef]
72. Dow, S.P.; Glassco, A.; Kass, J.; Schwarz, M.; Schwartz, D.L.; Klemmer, S.R. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Trans. Comput.-Hum. Interact.* **2010**, *17*, 1–24. [CrossRef]
73. Shah, J.J.; Smith, S.M.; Vargas-Hernandez, N. Metrics for measuring ideation effectiveness. *Des. Stud.* **2003**, *24*, 111–134. [CrossRef]
74. Hoßfeld, T.; Hirth, M.; Korshunov, P.; Hanhart, P.; Gardlo, B.; Keimel, C.; Timmerer, C. Survey of web-based crowdsourcing frameworks for subjective quality assessment. In Proceedings of the 2014 IEEE International Workshop on Multimedia Signal Processing, MMSP, Jakarta, Indonesia, 22–24 September 2014; pp. 22–24. [CrossRef]
75. Wu, H.; Corney, J.; Grant, M. Crowdsourcing Measures of Design Quality. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*; American Society of Mechanical Engineers: New York, NY, USA, 2014; Volume 46292, pp. 1–10. [CrossRef]
76. Woolley, A.W. Means vs. ends: Implications of process and outcome focus for team adaptation and performance. *Organ. Sci.* **2009**, *20*, 500–515. [CrossRef]
77. Dortheimer, J. A Crowdsourcing Method for Architecture—Towards a Collaborative and Participatory Architectural Design Praxis. Ph.D. Thesis, Tel Aviv University, Tel Aviv-Yafo, Israel, 2021.