

Article

A Qualitative-Quantitative Evaluation Model for Systematical Improving the Creativity of Students' Design Scheme

Lei Xiong ¹, Cheng-Lein Teng ², Yu-Qi Li ¹, Yuan-Zone Lee ², Bo-Wei Zhu ^{3,4,*} and Kun Liu ²

¹ Academy of Fine Arts, South China normal University, No. 55, Zhongshan avenue, Tianhe district, Guangzhou 510000, China; kmt20005@gmail.com (L.X.); liyuqi7718@163.com (Y.-Q.L.)

² College of Creative Design, Asia University, Lioufeng Rd. 500, Wufeng, Taichung 41354, Taiwan; ccd@asia.edu.tw (C.-L.T.); yrli1976@yeah.net (Y.-Z.L.); liukunphd@126.com (K.L.)

³ Faculty of Humanities and Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa 999078, Macau, China

⁴ Graduate Institute of Urban Planning, College of Public Affairs, National Taipei University, University Rd. 151, San Shia District, New Taipei City 23741, Taiwan

* Correspondence: bwzhu@must.edu.mo; Tel.: +853-65505064

Received: 30 March 2019; Accepted: 14 May 2019; Published: 16 May 2019



Abstract: For teaching architectural design, if creativity in students' practice assignments can be effectively evaluated, it will help make certain the ranking of the performance of different design schemes and provide a critical decision basis for the continuous improvement of each scheme. Based on the review of related literature, this study integrates the grounded theory (GT) from the social sciences' domain and the method of multiple attribute decision making (MADM) from operations research by analyzing their advantages and congruence to create a qualitative–quantitative evaluation model such as the GT-DANP-mV (GT, DEMATEL-based analytic network process [ANP] with modified VIKOR) model. Then, the characteristics and advantages of this evaluation model were examined by applying it to empirical research. Overall, this evaluation model fully utilized the advantages of the analytical techniques of GT and the MADM method to solve the corresponding problems at different stages of the evaluation study, which realizes their complementary advantages. More crucial is that, whether in the related evaluation study or practice assignments, the application of the GT-DANP-mV model could facilitate the formulation of continuous improvement strategies for architectural design schemes based on performance evaluation results under conditions of limited resources.

Keywords: evaluation model; design scheme; creativity; grounded theory; DANP-mV model (DEMATEL-based ANP with modified VIKOR model)

1. Introduction

Evaluative practices are important in all creative industries, where key individuals are invited to assess products “in-the-making” during initial and creative stages as well as to finalize products prior to communicating them to the market [1]. Nearly all design practices require some kind of evaluation or assessment [2]. The evaluation analysis is also an important part in design practice teaching, which takes the friendly and critical evaluation as its goal, with part of the assignments aiming at evaluating the latent characteristics, novelty, and market value of design products, as well as promoting creative explorations that strive toward innovation [1]. Evaluation of assignments provides designers with the opportunity to develop their designs, and, simultaneously, to become aware of their own design values, preferences, and sensitivity so that they come to reflect on their design processes and design conditions. Previous studies have indicated that “creativity” is one of the most important

performance indicators when evaluating a design [3,4]. For example, creativity is widely considered a cornerstone of architecture, and most, if not all, university programs refer to it as a desirable graduate attribute or intended learning outcome [5]. Although the assessment of creativity is considered an important issue in design education, research in the design domain has been and remains limited [4,6]. Perhaps, this is due to the elusiveness of the definition of the creativity of architecture students and professional designers as a professional attribute [7]. Malecki [8] considered that creativity in a very large measure remains unpredictable and hard to control. Therefore, the manifestation of creativity in the urban field can hardly be thoroughly managed and planned. Previous studies have shown that, in the psychological research domain, major obstacles to creativity studies are as follows: its mysterious origin, which are related to early studies, relatively independent of mainstream psychology, elusive conceptual definition, and uniform and limited research methods [9–11]. In this study, the target of evaluation is the “creativity” manifested in each architectural design scheme in the practice course, which is, by no means, evaluating the overall creativity of each student. In other words, the goals of the evaluation analysis lies in specifying the extent of “radical innovation” of the architectural design schemes in a practice environment with respect to solving practical problems and generating overall efficiency, and, in turn, providing a critical basis for ranking the performance of the practice assignments and the subsequent improvement of the students’ schemes.

From the perspective of cultural psychology, creativity is not a product of “disconnection,” but rather that of the ingrained “connections” between humans and environment, self and others, and creators and cultures [12]. The creativity of most people can be understood as social in nature, and because of this, those related studies that attempted to evaluate “creativity” were faced with the following problems: difficulty of comprehensively and systematically collecting related materials, difficulty of extracting evaluation criteria and judging their validity, and the evaluation analytical technique that can be used is still being too homogeneous [4,13]. However, this study clearly specifies the subjects of evaluation as the students’ schemes of an architectural design course, while the goal of evaluation is examining the “creativity” manifested in each design project. Even so, many previous studies have been criticized as overdependent on a single qualitative or quantitative evaluation [14]. Moreover, many were only content with identifying creative design projects rather than clearly specifying the varying degrees of creativity manifested in different schemes [3]. Furthermore, the evaluation methods applied in the evaluation studies in the architecture and other art design domains have often not given sufficient attention to training in criterion weights [1,4,6]. Overall, current evaluative analysis cannot practically promote the continuous improvement of students’ schemes in the process of teaching. Thus, there is an urgent need to develop a new evaluation model that can be applied to architectural design teaching for evaluating the creativity in practice assignments, to provide a decision basis for teachers and students to further improve their design schemes.

This study attempted to integrate cross-disciplinary research methods and analytical techniques for the construction of a reasonable and highly effective evaluation model for ranking the creativity performance in students’ practice assignments and further promote the continuous improvement of each design scheme. This study holds that the construction of this evaluation model requires the “embedment” of a mutually supporting process of dynamic thinking. Additionally, a systematic and mutual influence perspective is required to establish and confirm the influence of relationships and priority among each evaluation criterion. In this manner, the contributions of the evaluation study will not only be confined to the ranking and selection of design schemes but also provide a basis of reference for further formulating an improvement strategy [15,16]. Based on this, and along with the review of related literature, the various problems and difficulties that require a solution in evaluating the “creativity” manifested in the design schemes have been summed up. This study mainly focuses on the grounded theory (GT) from the social sciences domain and the multiple attribute decision making (MADM) from the operations research domain, which clarifies their different characteristics and advantages and disadvantages. The complementarity and congruence of both research methods will then be investigated, which is followed by the construction of a new evaluation model that

combines both qualitative and quantitative analytical techniques. Furthermore, it will demonstrate the research stages and procedures by applying this evaluation method to an empirical research. Lastly, the results of the research will be used to understand the positive impacts of this evaluation method on the sustainable development of design education.

2. Literature Review

2.1. Creativity in Design Schemes

Previous research has defined creativity in numerous ways. For instance, Amabile [17] defined creativity as “a production process that defines a certain (creative) object.” Torrance [18] described creativity as something “fluid, flexible, unique, and delicate.” Plucker et al. [11] believed that creativity is an attribute produced by an individual or a team under the interactive influence of three elements—ability, process, and environment. Sarkar and Chakrabarti proposed a “common” definition of creativity [19]: “Creativity occurs through a process by which an agent uses its ability to generate ideas, solutions, or products that are novel and valuable.” Since creativity is usually related to better design education and design practice [20], and is generally regarded as the prerequisite for invention and creation in design practice [3], design researchers have been inspired to conduct many empirical studies. In particular, some studies have focused on external inspiration to stimulate the ideation stage of the design process. These external inspirations, which are also considered tools that help designers obtain new ideas, are introduced into the design process, and, without these tools, such ideas will likely not emerge. Hence, many studies have stressed the positive results and merits of using external stimuli during the stage of ideation [21–23]. However, this might unfavorably interfere with the creation process when designers are providing sample solutions for the problems they are just considering. Such a phenomenon is called “design fixation” [24]. In studying this phenomenon, many researchers have extensively discussed the negative impacts of external stimuli [25,26].

Nevertheless, this study evaluates the creativity in design practice assignments. In the psychological research domain, many scholars have noted that creativity should not be treated as a cognitive variable in scientific research and assumed that it can be quantitatively identified [27,28]. This is because scholars generally consider that creativity emerges in complex social situations, which requires a more comprehensive and systematic method in understanding this concept. This is complex and hard to confine [10–12]. Yet, in this study, the concept of “creativity” can be understood as the extent of “radical innovation” of architectural design schemes in a practice environment with respect to solving practical problems and generating overall efficiency. Based on this, out of the consideration of solving practical problems, the use of critical quantitative analytical techniques at different stages of the evaluation study is indispensable. This is helpful in ranking the performance of different architectural design schemes and their selection and in formulating improvement strategies and action plans based on the evaluation results [29,30]. Many conventional quantitative analytical techniques operate on the assumption that factors are independent of one another. The widely used analytical hierarchy process (AHP) is included among them [15]. However, generally, important logical relationships exist among the various indicators (dimensions and criteria) for evaluating creativity in design schemes. In recent years, an increasing number of scholars have concentrated on analyzing the complex and hard-to-clarify mutual influence relationships among various evaluation indicators. Yet, this literature usually only investigated whether correlation existed between these various indicators instead of judging the direction and extent of the mutual influence among them, which would render the design schemes, having undergone an evaluation study, unable to practically make further improvement based on the evaluation results. The MADM method, which has been extensively used in the related studies in the domains of urban planning and design, has already been proven to be a practical and highly effective research method based on systematic thinking. It was combined with the perspective of the mutual influence among evaluation criteria, for dealing with the problem of “planning and evaluation” in the real world [31].

2.2. Multiple Attribute Decision Making (MADM)

As operations research has developed so far, the scholars within its domain has already developed many mixed MADM models, which have been widely used in different research areas [31–33]. First, conventional MADM assumes independent criteria with a hierarchical structure. However, relationships among criteria or dimensions are usually interdependent for real-world problems, and, in some cases, feedback effects exist. The criteria in practical MADM problems are generally interactive, and, thus, some inter-dependent models have been proposed (e.g., DANP, DEMATEL-based ANP). Second, conventional MADM only obtains relatively good solutions from existing alternatives, but also avoids “choosing the best among inferior alternatives,” i.e., avoids “Pick the best apple from a barrel of rotten apples,” it should be replaced by the aspiration levels [15]. Third, conventional MADM merely allows the selection and ranking of alternatives or strategies, but these alternative methods shift the focus from how to conduct “ranking” or “selection” of the most preferable alternatives to how to “improve” them [34]. Fourth, information aggregation, such as a fuzzy integral or a non-additive/super-additive model, has been developed for performance aggregation. Therefore, a hybrid dynamic (HD) MADM method is needed to overcome the defects of the conventional MADM method and solve the complications of dynamic problems in the real world [35]. In sum, this study considers that the advantage of the DANP-mV model in the MADM lies in its ability to clarify the influence relationships (i.e., influential network relation map, INRM) and priority (i.e., influential weights, IWs), among various evaluation criteria [36]. At the same time, the use of the modified VIKOR method can confirm the gap value between the performance of the empirical case and the desired standard, and, combining it with the INRM, it can precisely, and highly effectively, provide a critical basis for formulating continuous and systematic improvement strategies [15].

2.3. Grounded Theory

GT is a classic method used in qualitative research and was developed by two seasoned sociologists who are Barney Glaser from the University of Chicago and Anselm Strauss from Columbia University. Under the influence of interactionism and pragmatism scholars, they drafted the GT research method after studying multiple interactionism and pragmatism scholars as well as relevant research [37]. Strauss and Corbin [38] believed that GT is an approach used for examining and studying the real world. Since GT is designed to develop theories, a series of “methods” is used to collect and analyze data [39]. While observation, participant observation, field investigation, and interviews are the most common methods for collecting qualitative data, open coding, axial coding, and selective coding are the most common procedures used in the analytical process [40]. Sampling is used throughout the analytical process of the GT method. From collecting data and analyzing data to reaching the conclusion, all the steps are closely related [41]. Instead of developing a hypothetical theory before starting the research, researchers focus on developing theories (non-verification theories) during the research process without the influence of pre-existing theories. Data collection and analysis are done simultaneously. Following the formation of a new theory, researchers continue to collect the next batch of information. This process continues until theories derived from concepts achieve theoretical saturation [40].

Among all the qualitative research methods, GT is widely recognized as the most scientific method because it strictly follows scientific principles (e.g., inductive and deductive reasoning process), comparative principles, hypothesis testing, and the process of theory construction [42]. Thus, it is apparent that achieving theoretical saturation is the key point of this analytical process. Glaser and Strauss [41] pointed out that researchers who adopt GT should engage in continuous sampling until there is no more new or relevant data. They should also ensure that the development of different domains is comprehensive and that the links between the domains are well established and verified. Moreover, researchers should strike a balance between being innovative, cautious, consistent, and theoretically sensitive [39]. While theoretical sensitivity mainly comes from the professional and personal experience of researchers, it is common for researchers to be blinded from the truth by

pre-existing hypotheses, past experience, and literature during the analytical process. Therefore, researchers adopting GT must have a good grasp of the techniques designed to enhance theoretical sensitivity [40]. To sum up the above, the research process of GT is open-ended. The striking difference between this theory and other qualitative research methods is that the latter usually adopt a linear approach.

3. Establishing a New Evaluation Model by Integrating GT with MADM

In the process of an evaluation study, the primary task is generally to extract certain critical evaluation criteria from related information or data under the evaluation goals. In the course of architectural design teaching, students' practice assignments are usually simulating those design practice activities that provide solutions to some new social needs and phenomena in the real world. Much more, different kinds of sub-item design contents will be affected by the social situations in which an item is situated and face different practical problems. What is not difficult to infer is that the approach of solely relying on related literature to extract evaluation criteria does not suit this study. This would easily lead researchers to form a certain fixated system of understanding and knowledge structure with regard to the creativity manifested in architectural design schemes. Thus, this study considers that an effective qualitative analytical technique should be introduced in this evaluation model to deal with the related complex and massive qualitative data. It is worth noticing that the main characteristic of the qualitative analytical technique of GT is its emphasis on generating theories from the massive raw data collected, without relying on the old ones of previous studies. Researchers who use GT can gradually form their theoretical frameworks only by conducting in-depth analysis of the raw data [41]. Additionally, what is more critical is that GT provides a set of clear and systematic research procedures and analytical processes for researchers, which are significant characteristics compared to other qualitative research methods [43]. Previous studies have shown that this method emphasizes the logic of applied science, that is, induction, deduction, and hypothesis testing. These three stages are alternately, and circularly, operating in GT research [42,44].

On the other hand, regarding the problems of training in criterion weights and ranking the performance of design schemes, in light of the discussion above, those conventional quantitative analytical techniques operating on the assumption that factors being independent of one another are not suitable for this evaluation issue. This is because critical and crucial interaction effects often exist among the various criteria used to evaluate the performance of the design schemes. Furthermore, choosing to ignore the influence relationships among the evaluation criteria is not conducive to formulating corresponding improvement strategies for different cases based on the performance evaluation results [15]. Based on this, this study has noted that the major advantage of the DANP-mV model in MADM lies exactly in its ability to clarify the influence relationships (INRM) and priority (IW) among various evaluation criteria. Then, the application of the modified VIKOR technique to the performance evaluation analysis can confirm the gap value between the performance of the empirical case and the desired standard, and, combining it with INRM, it can precisely, and highly effectively, provide a critical basis for formulating supporting and continuous improvement strategies.

In sum, this study will construct an evaluation model according to different research stages and different qualitative and quantitative analytical techniques (as shown in Figure 1). Such a model in this study is called the GT-DANP-mV model. In the following subsections, the main research stages and concrete research procedures in the GT-DANP-mV model will be introduced in detail.

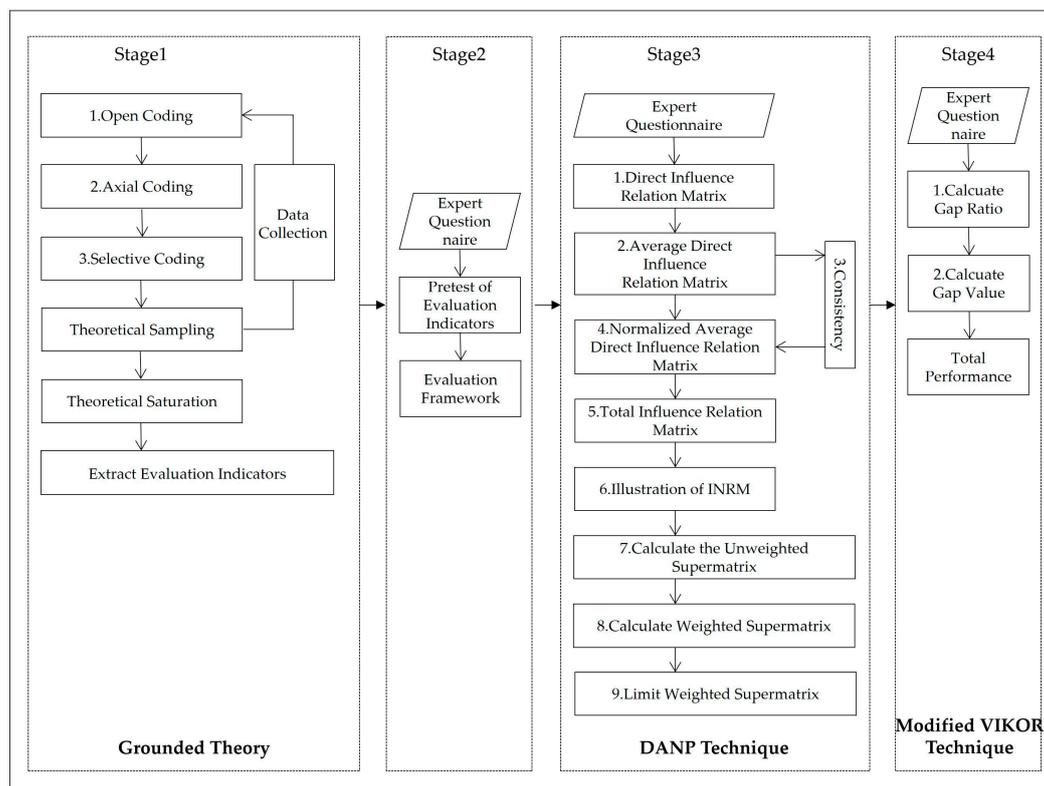


Figure 1. The GT-DANP-mV model.

3.1. Research Procedures of GT

In GT research, Barney Glaser and Anselm Strauss have developed a clear and systematic data analysis program for data coding. The coding mentioned in this section refers to the operation process of breaking up and conceptualizing the collected data, which is then followed by placing them together again in a new manner. Sampling will run through the analysis program of GT, with each procedure, from collection and analysis of data to ultimately obtaining the analysis results, intimately linked together. The most clear characteristic of GT is that both qualitative data collection and analysis take place simultaneously and alternately. After the first set of qualitative data are analyzed through the three procedures of open coding, axial coding, and selective coding, inductive and deductive methods are used to form a preliminary theory. This is followed by analyzing the second set of qualitative data to examine whether differences exist between the preliminary theoretical model and the analysis results of the second set. If differences exist, a comparison and modification of the analysis results of different data can be made to construct a more complete theoretical model. In the same manner, further data are collected through sampling until a “theoretical saturation” is reached. The three steps included in the data analysis program sequentially are:

- Open Coding

This step is the process of breaking up, examining, comparing, conceptualizing, and categorizing the data. Researchers can use questioning and comparative analysis in developing concepts. The collected data are first conceptualized, and then categorized and named. The process of categorizing is to combine a group of concepts into certain categories, which also include many subcategories.

- Axial Coding

This step mainly establishes the correlation between categories and subcategories and re-integrates the data to better develop the categories for creating a preliminary theoretical framework. Strauss and Corbin [45] suggested using a coding paradigm as the frame of reference for this analytical

procedure, which mainly includes elements such as antecedent conditions, theoretical phenomena, action/interaction strategies, consequences of action, context, and intervening conditions.

- Selective Coding

Following axial coding, it is necessary to go a step further to perform selective coding to establish a core category. A core category in general possesses “centrality.” The rest of the categories can revolve around it as the center to become closely interwoven, which, thus, integrates the theoretical framework of this research stage. In other words, the core category represents the main theme of a GT research. It highly condenses all the analysis results, with the aim to explain the main contents of this research stage.

3.2. Research Procedures of the DANP Technique

To untie the independence hypothesis, an increasing number of scholars in the operations research domain have started attempting to introduce the decision-making trial and evaluation laboratory technique (DEMATEL) into the MADM model. Moreover, by combining with the basic ideas of ANP and the foundation of allowing the existence of mutual influence relationships among various evaluation criteria, they also perform training in criterion weights. Therefore, this study will use the DANP technique to obtain the IWs of each evaluation criterion, clarify the complex influence relationships among all of them, and then draw an INRM. The concrete steps are shown below.

- Creating Direct Influence Relation Matrix E

The data were collected from the expert questionnaire, whose evaluation scores include 0, 1, 2, 3, and 4, with 0 representing the semantic meaning of having no influence and with 4 implying extremely influential. This questionnaire was filled out by experts of the research domain of this project, and the pairwise comparison method was used to evaluate the extent of influence of the criteria so as to express the extent of influence of one criterion on another under the evaluation framework. This matrix is an $n \times n$ -type non-negative matrix. Matrix E was created on the basis of the responses of several experts.

- Creating Average Direct Influence Relation Matrix A

The average matrix obtained by calculating the average score of several expert questionnaires was called the average direct influence relationship matrix.

- Consensus Testing of Expert Questionnaires

The consensus testing of the expert questionnaires was performed by calculating the average gap ratio. The threshold value of the average gap ratio in statistics is 5%. When the calculation result is smaller than 5%, it means that the confidence level of the collected expert questionnaires is more than 95%. It can be regarded as a stable and reliable system, and these questionnaires have manifested a relatively high consensus. However, if the testing results show that it is not a stable system, it is then necessary to re-examine whether the collected data are correct or whether the number of experts is adequate.

- Creating the Normalized Average Direct Influence Relation Matrix D

The result of normalization transformed all the numerical values in matrix A correspondingly into the numerical values between 0 and 1. In the matrix, all the numerical values located at the diagonal are 0.

- Creating Total Influence Relation Matrix T

When the powerful number of matrix D approximates to infinity, the indirect influence in it would continuously decline. Thus, it is first necessary to create an inverse matrix of the difference between an $n \times n$ identity matrix I and matrix D . By multiplying this inverse matrix and matrix D , the total influence relation matrix T could be obtained.

- Drawing INRM

This step requires quantifying the influence effects each criterion produces and is received by other criteria. This could be obtained by adding up the numerical value of each line and each row in matrix T , and, at the same time, a total dimensional influence relation matrix T_D was constructed. Then, the INRM for various dimensions and various criteria were drawn respectively.

- Creating Unweighted Super Matrix W^u

First, a normalization calculation was performed on the total criterion influence relationships according to different dimensions. Then, based on the ANP concept of pairwise comparison of criteria, and, through conversion, the normalized influence relationship matrix T_C^u was converted to the unweighted super matrix W^u .

- Creating the Weighted Super Matrix

First, the numerical value of each line in the total influence relationships matrix T_D were added up to obtain d_i , which then produced matrix T_D^u through normalization. Lastly, by multiplying the unweighted super matrix W^u and matrix T_D^u , the weighted super matrix W was obtained.

- Calculating the Limiting Super Matrix W^s

Weighted super matrix W , through self-multiplication, would eventually end up in the convergent and stationary matrix W^s . Having undergone limiting, the numerical value in each line of matrix W^s would be global weight w^s of the criteria it corresponded to, which is also referred to as IWs.

3.3. Research Procedures of the Modified VIKOR Technique

This study will use the modified VIKOR technique to handle the collected performance questionnaires. This analytical technique can clearly specify the gap values between the performance of the empirical case study and the desired standard in the existing conditions. At the same time, it also improves the VIKOR technique's setting of a positive-ideal solution and a negative-ideal solution. The concrete steps are as follows.

- The modified approach for replacement by the aspiration level and the worst value is as follows.

The aspiration level: $f^{aspired} = (f_1^{aspired}, \dots, f_j^{aspired}, \dots, f_n^{aspired})$, where $f_j^{aspired}$ is an aspiration level, or is called the best value.

The worst values: $f^{worst} = (f_1^{worst}, \dots, f_j^{worst}, \dots, f_n^{worst})$, where f_j^{worst} is a worst value.

In this study, performance scores ranging from 0 to 10 (very bad \leftarrow 0, 1, 2, ..., 9, 10 \rightarrow very good) are used with natural language in the linguistic/semantic questionnaire. Thus, the aspiration level takes the highest score of 10 and the worst value takes the value of 0. Hence, $f_j^{aspired} = 10$ is defined as the aspiration level and $f_j^{worst} = 0$ as the worst value.

- Determine the mean group utility for the gap and then establish the priority improvement strategy. These values can be calculated using Equation (R1) below.

$$s_k = \sum_{j=1}^n w_j r_{kj} = \sum_{j=1}^n w_j (|f_j^{aspired} - f_{kj}|) / (|f_j^{aspired} - f_j^{worst}|) \tag{R1}$$

where s_k is defined as the normalized ratio (%) of distance to the aspiration level, which implies the synthesized gap of the criteria. In this case, w_j indicates the IWs for the criteria obtained from tDANP.

4. Empirical Case Study

4.1. Empirical Case Description

The course of this empirical research is from the Faculty of Humanities and Arts, Macao University of Science and Technology. As a new higher education institution in East Asia, the university has been expanding in recent years. The faculty offers two specialties: (1) Communication Studies and (2) Design Studies. The Department of Environmental Design in the faculty has about 60 students, seven full-time teachers, and four part-time teachers. The architectural design course discussed in the empirical case study is a mandatory course in the undergraduate program. Its pedagogical task is allocated to three full-time teachers, who all hold a doctorate in architecture and urban planning. Each person is in charge of a class, with an average of 27 junior undergraduates each. The purpose of this course is to help students master the principles and basic processes of the design of small public architectures and understand the general architectural requirements and relevant specifications of cultural exhibition halls. The architectural design task in this course was not a virtual subject but a new exhibition hall in a designated location, which will be used to display the teaching and scientific research achievements of Macao University of Science and Technology. The design should comply with the construction plan of the campus and the development requirements of the Faculty of Humanities and Arts. The site is located at the junction of Avenida do aeroporto and Avenida Dr. Henry Fok, behind block P of the dormitory building of Macao University of Science and Technology, which is in the Aberdeen district of the Macao Special Administrative Region (Figure 2). The surrounding buildings include the main building of the University International College, dormitory building, gymnasium, and golden carpentry laboratory building of Macao University of Science and Technology. The site area of the project is about five hectares. The construction area should be between 1500 m² and 2500 m² and the maximum stories of the building are three. Since the students were only required to complete a conceptual design scheme within this architectural design course, this project finished in only seven weeks. However, after the end of the course, the university will provide full support to teachers and students to spontaneously and continuously develop such creative design schemes. They will then participate in the project's publicity and tender activities organized by the university.

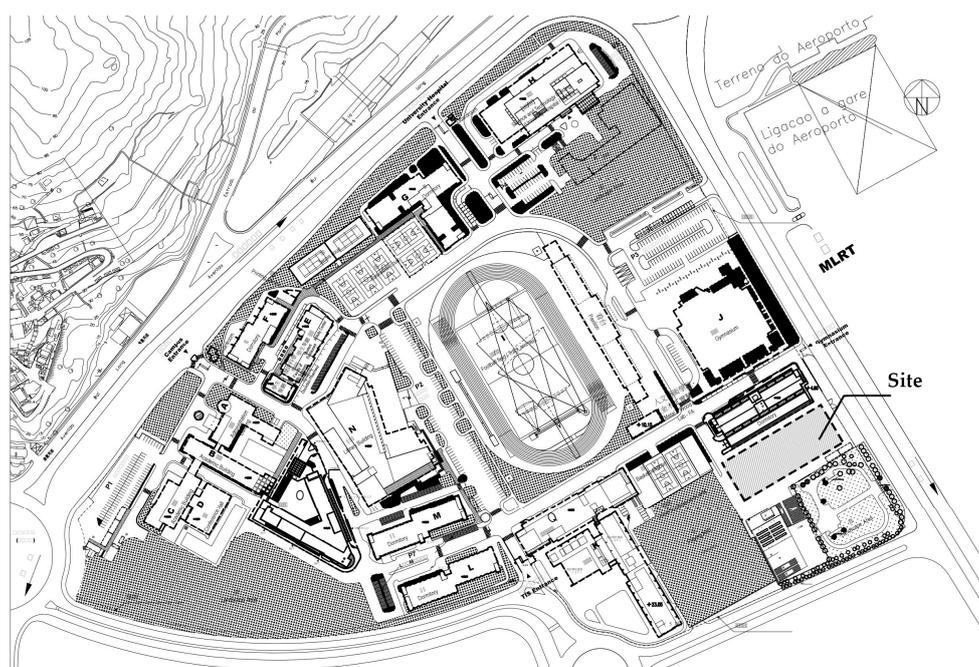


Figure 2. The site of the exhibition hall of MUST.

4.2. Documents and Data Sampling

Based on the discussion above, the GT-DANP-mV model can be divided into three main stages of information and data collection according to the different analytical techniques used at different research stages. First, this study conducts a GT research to extract the indicators (dimensions/criterion) that can be used to construct the assessment framework. At this stage, this study collects qualitative data through expert interviews. Second, to clarify the influence relationships and priority of the evaluation indicators, this study collects relevant data by issuing expert questionnaires and applies the DANP analysis technique to complete this research stage. Lastly, we issue the questionnaires again to collect relevant data and apply a modified VIKOR analysis technique to complete the performance ranking of empirical cases. Based on this, we will present an in-depth introduction regarding the document analysis and data sampling process in the following three subsections.

4.2.1. Participants and Interviews

The idea behind the data sampling in GT is to select participants who will best contribute to the understanding of the problem and the research question. Therefore, this study will focus on in-depth interviews with two types of experts at this stage. One type of expert group comprises full-time teachers from the Department of Architectural Design of the Faculty of Humanities and Arts in Macao University of Science and Technology, all of whom hold a doctorate and most of whom have been engaged in related research. All the interviewees have more than five years of teaching and research experience and have served as principal lecturers in architectural design courses many times. The other group of experts comprises senior designers from the architectural design industry. As respondents, designers from reputable design institutes have acquired enough basic information to understand the case in this empirical study. Some have more than five years of work experience while others hold at least a master's degree in architecture or related disciplines. All interviews were conducted in their personal studio from 24 February 2018 to 20 May 2018, and the length of each interview ranged from 30 to 120 minutes, with an average of 60 minutes. In the course of the interview, questions were raised according to the purpose of different research stages, such as the following.

- "How do you understand creativity in architectural design practice?"
- "According to your personal teaching experience, what breakthroughs can students often achieve in conceptual designs that can demonstrate creative thinking?"
- "Have any of your students ever impressed you with their unique creative thinking through a design scheme in a similar course, and how was it elaborated?"
- "As for the exhibition hall of Macao University of Science and Technology, what realistic problems can be solved with creative thinking?"

All interview questions are semi-structured, and the author adjusted the order of the questions, some details, and even some questions accordingly. One element of GT is the practice of concurrent data collection and analysis that subsequently guides further participant selection and is known as theoretical sampling [41]. After collecting and analyzing the data, this study determined the next interviewee according to the theoretical concept. The work of sampling continued until the information in one category reached "theoretical saturation." This process emphasizes the importance and richness of the data but not its quantity [40].

4.2.2. DANP Questionnaires

In this study, we want to investigate, through the second stage of research, how architectural designs can be made more creative in empirical research. Therefore, the participants of the DANP questionnaire must be experts who have worked in the architectural design industry or conducted teaching research. Since the experts had provided a thorough and comprehensive perspective based on their own experience and knowledge and the evaluation indicators (dimensions/criterion)

were collected from qualitative data, the DANP questionnaire was still issued to experts who have participated in the previous interview, which enables these key experts to continue pondering this research and to provide more in-depth perspectives. Lastly, the DANP questionnaire was conducted over a period of one month, and seven questionnaires were distributed with seven valid and zero invalid questionnaires. After GT research, the respondents at this stage were already determined, and personal information regarding these experts is provided in detail in Table 1. The seven experts were instructed on the research questions and objectives and how to respond to the questions. The experts completed the questionnaire after the note questionnaire. To improve validity, questionnaires were administered in the form of structured interviews, during which any confusion about the questions was clarified by the authors. The average time for questionnaire completion was approximately 2.5 hours.

Table 1. Participants' information.

| Experts | Years Worked | Job Title | Academic Degree | Sex |
|---------|--------------|---------------------|------------------------|--------|
| A | 15 | Associate Professor | Doctor of Architecture | Male |
| B | 10 | Associate Professor | Doctor of Architecture | Male |
| C | 5 | Assistant Professor | Doctor of Architecture | Female |
| D | 5 | Assistant Professor | Doctor of Design | Male |
| E | 8 | Architect | Master of Architecture | Female |
| F | 10 | Architect | Master of Urban Design | Male |
| G | 7 | Architect | Master of Urban Design | Male |

4.2.3. Modified VIKOR Questionnaire

After establishing the assessment framework and determining the relationships and priority of different assessment indicators (dimensions/criterion), this study used modified VIKOR questionnaires to evaluate empirical performance. Experts who participated in the first two stages were still involved, and questionnaires were given face-to-face to each expert in their personal studio. In this empirical study, there were nine sets of architectural designs to be evaluated. Before completing the modified VIKOR questionnaire, each respondent has had a complete and continuous understanding of all the designs, and the design drawings were presented to them while they were filling out the questionnaire. The study lasted for 15 days, and seven questionnaires were issued. None of the questionnaires were invalid. The average answering time for each questionnaire was one hour. The author will further discuss all the problems that occurred during the process.

4.3. Results and Discussion

4.3.1. Extraction of Evaluation Indicators Based on GT

Building on Strauss and Corbin [39], we first analyzed the data with an open coding system, followed by an axial coding system. All interviews were transcribed and analyzed before further data collection occurred. Data were manually coded using line-by-line analysis. In accordance with the above steps, the data were broken down into discrete ideas, events, and acts. Each phenomenon was related to the participants' attitudes toward the creativity of this architectural design project in empirical case studies. The open coding process enabled the development of themes that were grounded in the data itself. Meanwhile, during the process of constant comparative analysis, within and between data sets, the concepts were examined for similarities and differences. Data collection and analysis continued until all concepts clearly explained what was happening within the field of study. Lastly, concepts were examined for the relationships. These concepts were then developed into categories, which explain how to evaluate the creativity of the design project of the Macau University of Science and Technology Exhibition Hall. Memo writing to record analytic thinking, theoretical development, and methodological decisions were written after each interview and were further developed as concepts during ongoing data collection and analysis (as shown in Appendix A, Table A1).

Table A1 shows the process of open, axial, and selective coding based on GT, in which the process involved sorting memos, labeling data, conceptualizing data, categorizing data, and subcategorizing data. After the process of conceptualization, two items (a1 and a5) can be sorted out into one (aa1), which concerns site planning. The total labeled number was 115. After conceptualizing, the number decreased to 83. To a certain extent, the concepts became clearer through open coding. Axial coding occurred concurrently with open coding. In the data analysis step, the data were broken down into concepts and categories and then reassembled by axial coding. In the process of coding, a concept was defined whenever two or more respondents mentioned a certain phenomenon. Furthermore, 83 initial concepts were unstructured. The next step was to gather these concepts into categories based on their similarities or differences. For example, in the process of conceptualizing data, the concepts aa1 and aa9 were gathered into category A1. Therefore, 46 categorized items were generated (Table A1). When coding axially, the main point was to determine how categories link and crosscut. For example, the relationship among A1, A4, A7, A10, A15, and A18 were similar, so they were also allocated to one category labeled architectural functionality (AA1). Once developed categories emerged, selective coding was started. Central to the procedure of selective coding was the selection of a core category, major categories related to it, and to one another.

Through the above analysis procedures, three major categories were generated, and there existed an obvious hierarchical relationship between categories and subcategories (as shown in Figure 3). The mutual influence between subcategories under each category was involved in the interview process, but there existed no superiority–inferiority relationship between the abstraction levels of the three major categories. In other words, there existed no inclusive relationship between major categories. Results of the analysis show that the evaluation of creativity contained in this design practice is a multi-dimensional and hierarchical thinking process. From different perspectives, the degree of creativity associated with the corresponding category can be divided into three types: architectural functionality (Logical Reasoning), architectural artistry (Aesthetic Experience), and architectural sociality (Impression Analysis). In sum, the core category of the GT research stage in this empirical study is the evaluation mechanism of the creativity of the design practice. The core category was highly condensed above the analysis results, and other categories can be closely combined around this core category.

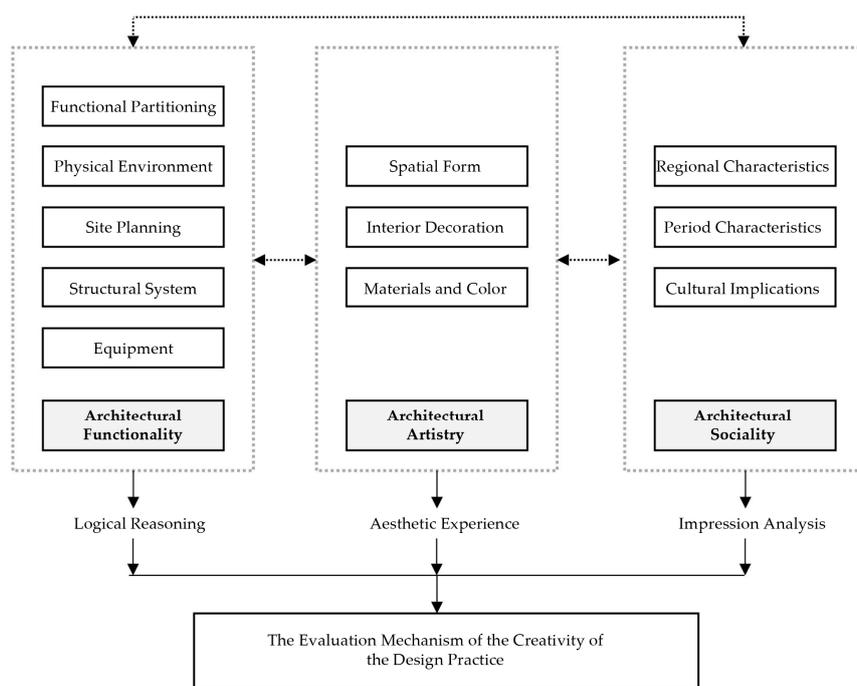


Figure 3. Analysis of the evaluation mechanism.

4.3.2. Pretest of Evaluation Indicators

The evaluation criteria extracted in the GT research stage needed to be pretested first. Evaluation analysis in the next stage was conducted within the clearly defined evaluation framework only after the effectiveness of evaluation criteria was confirmed. According to the pretest analysis procedure in the conventional evaluation research, the pretest stage of this study includes two phases of three steps. In the pretest stage, however, the respondents were all among the interviewed experts in the previously mentioned GT research stage. Therefore, the respondents verified the effectiveness of the evaluation criteria extracted through the GT research based on the previous verbatim transcription and the author's coding process and results.

The first phase of the pretest stage was to survey the experts' comments on the verbatim transcription and coding analysis results through semi-structured questionnaires. The experts first checked the consistency between the verbatim transcription and the respondents' narrations, and then checked the analysis results against the experience of the respondents. If surveyed experts asked for clarifications or noted ambiguities in the analysis results, this phase was repeated. Through questionnaire surveys, the author collected the comments and judgments on controversial issues identified by all experts surveyed during the pretest stage. If the issues were clarified, the author modified the established evaluation framework according to the experts' comments. Upon completion of the first phase, it was basically confirmed that the evaluation framework was stable and possessed sufficient reliability.

In the second phase of the pretest stage, the degree of importance of each evaluation criterion was tested through a structured expert questionnaire. Specifically, according to the collected data, the degree of importance of each evaluation criterion was determined by the average index. During the process, the experts gave a score (0 to 10) to the degree of importance of each evaluation criterion, according to their experience and cognitive knowledge. Specifically, 0 means "extremely unimportant" and 10 means "extremely important." The author set the score of 7.5 as the threshold, which indicates that, if the score of an evaluation criterion was higher than 7.5, the degree of importance reached a level of 75% or above. In this sense, the experts believed that the evaluation criterion in the evaluation framework was important, and, thus, should be retained. For evaluation criteria with a score of 5 to 7.5, an expert interview was held to discuss why they were defective, whether they needed to be retained, and how they needed to be modified. The author integrated the data in the expert survey, and then a questionnaire survey was conducted again to seek the experts' comments on the modified evaluation framework. This process was performed repeatedly until the score of each evaluation criterion was not lower than 7.5. Upon completion of the pretest stage, the evaluation framework used in this study was confirmed.

In this empirical study, seven experts attended the pretest of evaluation criteria. Upon completion of the first phase of the pretest stage, no experts raised any objections concerning the verbatim transcription recorded by the author in the GT research stage, and the surveyed experts all agreed to the previous analysis results. Evidently, the evaluation framework built in this study possessed sufficient reliability and was essentially a stable evaluation system. However, this did not suffice to prove that these criteria were effective regarding the topic in this study. Therefore, it was necessary to perform the second phase of the pretest stage, that is, to conduct a further test of the degree of importance of each evaluation criterion. Noteworthy, the "degree of importance" in this phase indicated whether an evaluation criterion was important enough to be retained in the evaluation framework while experts evaluated the creativity of the reformation design of the idle space in the evaluated creative community. As mentioned above, the threshold in this phase was set to 7.5. After the second phase of the pretest stage, this study did not find any evaluation criterion with a score lower than 7.5. Therefore, all evaluation criteria extracted by the GT research were important and effective, and no modifications or deletions were required. In sum, upon completion of the pretest stage, the evaluation framework built in this study comprised three evaluation dimensions and 11 evaluation criteria (corresponding to the categories and subcategories in Figure 3).

4.3.3. Applying the DANP Technique to Distribute IWs and Drawing INRM

As set forth above, the three categories in this research stage were viewed as dimensions of the evaluation framework and the 11 subcategories were viewed as evaluation criteria. The intent of this research stage was to assign IWs to all evaluation indicators (dimensions and criteria), draw an INRM, and, thus, determine which indicators were dominant, to determine the complex influential relations in this evaluation system (as shown in Figure 4). Figure 4 includes four INRMs. The influential network relations were drawn according to the total influence relationship matrix created by this study (as shown in Appendix A, Table A2). The horizontal axis in the figure is the “centrality ($r+s$)” of each criterion in that system, while the vertical axis is the “relatedness ($r-s$)” of each criterion. Table A3 in the Appendix A shows the calculating process of the centrality and relatedness of each evaluation criterion. Since the relatedness of the dimension of the sociality of architecture (D_3) in this evaluation framework was the greatest, this implies that the D_3 dimension occupied the dominating influential position. Its influence on the other two dimensions was stronger than what was exerted on it. By comparison, the aesthetics of architecture (D_2) was the dimension with the lowest dominating influence effect. Besides the influence relationships among each dimension, Figure 4 also clearly demonstrates the influence relationships among each evaluation criterion under each dimension. The D_1 dimension comprised five evaluation criteria. Specifically, site planning (C_{13}) had the strongest influence, equipment (C_{15}) had the second strongest influence, and the influence of the three other criteria was ranked in the following order: $C_{14} > C_{12} > C_{11}$. The D_2 dimension comprised three evaluation criteria. Specifically, spatial form (C_{21}) played a dominant role, material and color (C_{23}) had the second strongest influence, and the interior decoration (C_{22}) had the weakest influence. The D_3 dimension comprised three evaluation criteria. Specifically, cultural implication (C_{33}) had the strongest influence, and the period characteristic (C_{32}) had the second strongest influence.

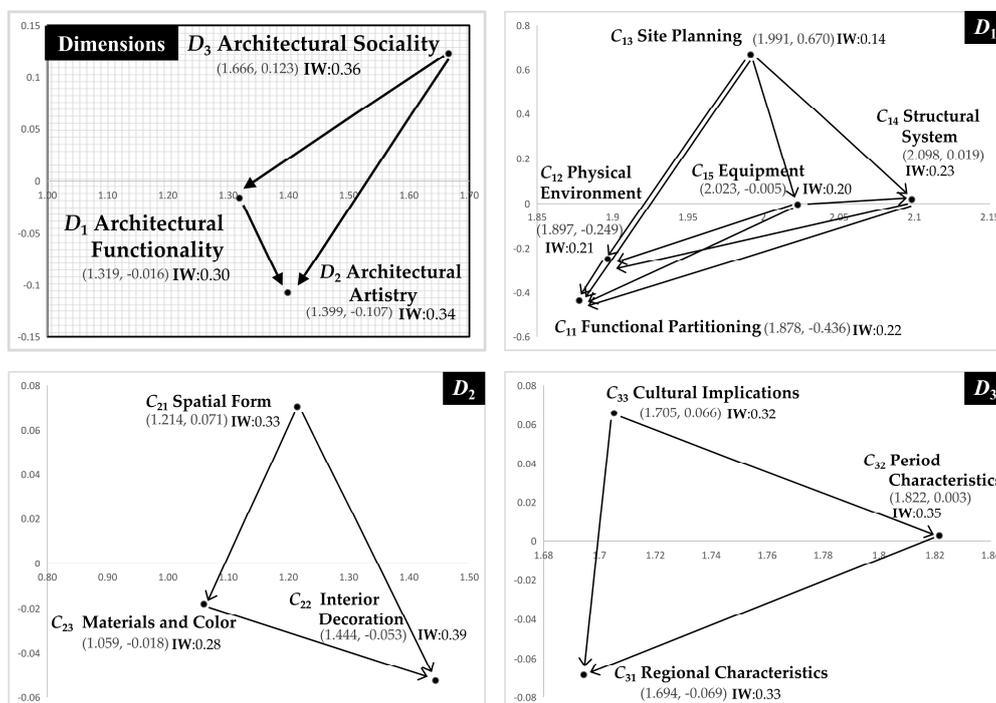


Figure 4. The INRM (influential network relation map) of total influence relationships.

As shown in Figure 4, architectural sociality (D_3) is the dimension with the highest IW in the evaluation framework. Meanwhile, the three dimensions in the evaluation framework are of approximate IW. It can be said that the three dimensions are all important evaluation indices. Under the current research topic, the dimension that has the highest IW is dominant. Using the DANP

technique, all evaluation indicators are assigned their IWs, which facilitates the subsequent analysis of performance evaluation in this study. Furthermore, determining the INRM and IWs between/of the evaluation indicators of the current research topic will play a very critical role in promoting the continuous and systematic improvement and update of the architectural design practice.

4.3.4. Applying the Modified VIKOR Technique to Conduct a Performance Evaluation Analysis of the Empirical Case Study

The last stage of this empirical study was to conduct a performance evaluation for course assignments using modified VIKOR technology. As mentioned above, this architectural design course demanded that nine groups of students complete course assignments (as shown in Figure 5). Figure 5 presents each group of design schemes through representative drawings. While filling in the modified VIKOR questionnaires, the experts could view all detailed drawings of each group of design schemes, so that the creativity of the design schemes was fully exhibited in the presence of the experts. Performance evaluation showed that the No. 8 design scheme was of the highest creativity while the No. 4 design scheme was of the lowest creativity. By degree of creativity, all design schemes were ranked in descending order as follows: No. 8, No. 6, No. 7, No. 5, No. 9, No. 2, No. 1, No. 3, and No. 4 (as described in Table 2).

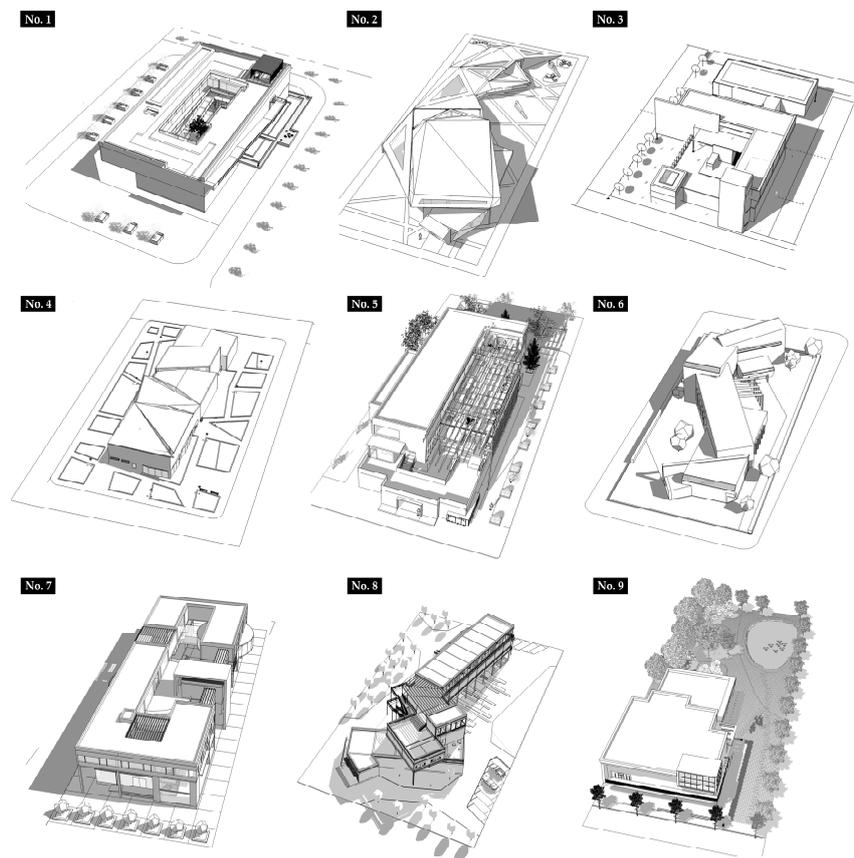


Figure 5. Architectural design schemes of the MUST exhibition hall.

Table 2. The performance evaluation of the architectural design schemes.

| Dimensions/Criteria | Influential Weights (IW _s) | No.1 | No.2 | No.3 | No.4 | No.5 | No.6 | No.7 | No.8 | No.9 |
|--------------------------|--|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | | Gap Ratio |
| <i>D</i> ₁ | 0.30 | 0.68 | 0.48 | 0.57 | 0.77 | 0.29 | 0.41 | 0.37 | 0.30 | 0.56 |
| <i>C</i> ₁₁ | 0.22 | 0.49 | 0.30 | 0.27 | 0.76 | 0.44 | 0.39 | 0.19 | 0.14 | 0.43 |
| <i>C</i> ₁₂ | 0.21 | 0.71 | 0.31 | 0.63 | 0.91 | 0.26 | 0.46 | 0.47 | 0.17 | 0.51 |
| <i>C</i> ₁₃ | 0.14 | 0.49 | 0.51 | 0.76 | 0.93 | 0.31 | 0.33 | 0.24 | 0.31 | 0.26 |
| <i>C</i> ₁₄ | 0.23 | 0.83 | 0.86 | 0.94 | 0.96 | 0.29 | 0.60 | 0.56 | 0.63 | 0.73 |
| <i>C</i> ₁₅ | 0.20 | 0.81 | 0.40 | 0.29 | 0.29 | 0.16 | 0.21 | 0.36 | 0.21 | 0.79 |
| <i>D</i> ₂ | 0.34 | 0.71 | 0.67 | 0.72 | 0.85 | 0.71 | 0.58 | 0.43 | 0.38 | 0.43 |
| <i>C</i> ₂₁ | 0.33 | 0.59 | 0.31 | 0.61 | 0.69 | 0.76 | 0.19 | 0.34 | 0.16 | 0.23 |
| <i>C</i> ₂₂ | 0.39 | 0.79 | 0.84 | 0.93 | 0.97 | 0.54 | 0.71 | 0.41 | 0.43 | 0.47 |
| <i>C</i> ₂₃ | 0.28 | 0.76 | 0.84 | 0.54 | 0.86 | 0.86 | 0.84 | 0.54 | 0.59 | 0.61 |
| <i>D</i> ₃ | 0.36 | 0.82 | 0.78 | 0.87 | 0.94 | 0.60 | 0.45 | 0.67 | 0.64 | 0.75 |
| <i>C</i> ₃₁ | 0.33 | 0.87 | 0.93 | 0.96 | 0.99 | 0.24 | 0.17 | 0.34 | 0.51 | 0.70 |
| <i>C</i> ₃₂ | 0.35 | 0.73 | 0.53 | 0.70 | 0.91 | 0.81 | 0.51 | 0.86 | 0.64 | 0.73 |
| <i>C</i> ₃₃ | 0.32 | 0.87 | 0.90 | 0.96 | 0.93 | 0.74 | 0.66 | 0.81 | 0.76 | 0.83 |
| Total Performance | | 2.59 | 3.48 | 2.73 | 1.44 | 4.57 | 5.21 | 5.00 | 5.51 | 4.14 |

The performance evaluation showed that the disadvantages of the No. 4 design scheme mainly lay in architectural sociality (*D*₃) and architectural artistry (*D*₂). That is, the experts did not feel any creativity in the artistic expression of this design scheme. Moreover, this design scheme did not express any cultural implications. The advantages of the No. 8 design scheme mainly lay in the *D*₁ and *D*₂ dimensions. The surveyed experts universally believed that the architectural form in this design scheme was full of novelty and aesthetic beauty. Meanwhile, this design scheme considered various advantages and disadvantages of the site conditions, which produced an original response regarding various restrictions, and addressed the restrictions properly through a reasonable design.

According to the previously mentioned study results, the modified VIKOR technology applied to this evaluation pattern has the following main advantage: preventing the defect of the traditional performance evaluation pattern, known as “selecting a less rotten apple from many rotten apples.” In other words, performance ranking in the traditional performance evaluation pattern is usually based on the relative evaluation results (for example, excellent, good, and bad) of evaluated objects. As a result, the object with the best performance among all evaluation objects is usually given full cumulative scores with respect to one or several evaluation indicators. This implies that the object needs almost no improvement or has no space for improvement. In contrast, the object with the worst performance is usually given a zero cumulative score with respect to most evaluation indicators and does not have good performance or advantages. Clearly, this is not the case in the real world. Furthermore, the gap values obtained by the modified VIKOR technology can be combined with the INRM drawn in the previous study stage, which helps us develop a continuous and systematic improvement strategy suited to empirical cases [29,30]. For example, for the No. 5 design scheme with an average overall performance, the *D*₂ dimension is given the lowest cumulative score. If various efforts and costs were continuously inputted during the subsequent modification process to improve architectural artistry, this may fail to improve the overall performance of this design scheme genuinely and effectively. Based on the previously mentioned INRM (as shown in Figure 4), this study recommends that the designers focus on architectural sociality, or to be specific, increase the period characteristics presented by the design scheme, and develop a series of chain effects to exhibit its architectural artistry better and improve its overall performance dynamically and continuously. As set forth above, this will fundamentally overcome the existing defects of the known traditional performance evaluation pattern. This means to take stopgap measures only but ignore the essential solution.

5. Conclusions

This study finds that the GT-DANP-mV evaluation pattern can ensure all evaluation indicators are observed and analyzed in the same system during the process of creative evaluation. In previous studies, certain scholars evaluated the creativity of design schemes in terms of different evaluation dimensions and on a stage-by-stage basis [3]. This is beneficial when discussing the complex and critical influential relationships between evaluation indicators more deeply (regardless of whether the evaluation indicators fall under the same dimension). In traditional performance evaluation studies, evaluation indicators are extracted through literature analysis. The GT-DANP-mV evaluation pattern can make the extracted evaluation indicators more significantly target-oriented (effective). It can even be said that the categories and subcategories obtained by the GT research are totally based on the original qualitative data and are exclusively used for this empirical study. The pretest in this evaluation pattern is conducted subsequent to the GT research, so it is able to select experts for the pretest stage and a subsequent performance evaluation phase more accurately, according to the early interview results. Meanwhile, the discussion in the pretest stage may focus on the verbatim transcription and coding results. Compared with the pretest phase in the traditional performance evaluation pattern, the GT-DANP-mV evaluation pattern allows experts to judge the effectiveness of evaluation criteria without recourse to any reference data. This empirical study also shows that qualitative and quantitative study stages contained in the GT-DANP-mV evaluation pattern do not overlap or conflict with each other. By combining the advantages of two analysis techniques (including GT and MADM), the GT-DANP-mV evaluation pattern is also capable of solving the problems encountered at different stages of the performance evaluation. It is not difficult to find that, after DANP technology is applied subsequently to GT research, analysis results of selective coding can be verified to a certain extent, which further extends and supplements the subsequent analysis. Moreover, the introduced DANP-mV hybrid multi-attribute decision model is able to implant the sustainability philosophy into the study on the evaluation of the creativity of an architectural design practice. The DANP-mV model can genuinely carry out the philosophy of continuous improvement and sustainable updates, complete evaluation analysis, and performance ranking in a dynamic, interactional way of thinking, and, thus, develop a continuous and systematic improvement strategy. This will enable the students' design schemes to be continuously updated in the process toward a desired level (creativity). In the follow-up study, a non-additive performance analysis technique may be used tentatively instead of the modified VIKOR technique used in this study. Therefore, the modified VIKOR method used in this study to assess performance levels in the case studies is an additive method. However, circumstances in practice may often be non-additive, and, thus, follow-up research may use non-additive methods to assess performance by more closely approximating actual circumstances.

Author Contributions: L.X. and B.-W.Z. designed the research. Y.-Z.L. and C.-L.T. provided the research idea. L.X. and B.-W.Z. collected the data, analyzed the data and drafted the manuscript. Lastly, L.X., Y.-Q.L. and K.L. revised the paper.

Funding: This study is funded by the Faculty Research Grant (FRG) of Macau university of Science and Technology, and the project number is FRG-18-037-FA.

Acknowledgments: The authors are extremely grateful for the sustainability journal editorial team's valuable comments on improving the quality of this article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Results in Detail

Table A1. Coding process for open coding, axial coding, and selective coding based on GT.

| Sorting Memos | Labeling | Conceptualizing Data | Categorizing Data | Categories | Subcategorized |
|---|---|--|--|---|--|
| (How do you understand creativity in architectural design practice?) “Creativity in this work is usually embodied by solving actual difficulties or problems.” “Regarding the architectural design scheme, creativity implies not only novelty, but also efficiency and practicality.” ... | a1 Highlight the harmonious integration of the service infrastructure with the greenery landscape at the site and the building as a whole. a2 Variation and unity of colors in the internal/external spaces of the building. a3 Convey the design philosophy through reasonable selection, matching, and treatment of materials of the building façade. a4 Add flexibility and variability to the space, to satisfy the diverse perceptual needs of the users. a5 Fully explore the correlation between the site conditions and the architectural layout. a6 The architectural appearance should possess a certain aesthetic beauty and visual appeal. a7 Reflect social issues through the architecture. | aa1 Combine various elements of the site and the building into an organic whole and intensify the correlation between the building and the site. (a1 and a5) aa2 Matching of colors in the internal/external spaces of the building and coordination with the campus environment (a2 and a17) aa3. Convey the design philosophy through reasonable selection, matching, and treatment of materials of the building façade (a3) aa4. Add flexibility and variability to the space, to satisfy the diverse perceptual needs of the users (a4) aa5. The architectural appearance should possess a certain aesthetic beauty and visual appeal. (a6) d6. Reflect local social development through the architectural design (a7 and a10). | A1 The building and its site should constitute an organic whole, and improper development of site resources should be prevented to the greatest extent. (aa1 and aa9). A2 Architectural color design that is varied and harmonious with the overall environment (aa2 and aa31) A3. The façade materials should respect the philosophy of sustainability (for example, energy conservation and environmental friendliness), and the integrity of the building façade should be ensured. (aa3 and aa19) A4. The architectural space should satisfy the diverse and personalized needs of the users, and the “uncertainty” of architectural space should be intensified. (aa4 and aa15) A5. Create unique and aesthetically pleasing internal/external spatial forms for the building (aa5 and aa13). A6. Attempt to convey new knowledge and understanding of the architectural ontology, and intensify the correlation between the building and local socio-economic development, and symbolic meaning of the building (aa6 and aa20). | AA1 Architectural Functionality AA2 Architectural Artistry AA3 Architectural Sociality | AA1 Architectural Functionality c11 Functional Partitioning c12 Physical Environment c13 Site Planning c14 Structural System c15 Equipment AA2 Architectural Artistry c21 Spatial Form c22 Interior Decoration c23 Materials and Color AA3 Architectural Sociality c31 Regional Characteristics c32 Period Characteristics c33 Cultural Implications |
| Initial data collection | 115 items | 83 items | 46 items | 3 items | Generating the substantive theory |

Table A2. Construct the total influence relation matrix T.

| T | C ₁₁ | C ₁₂ | C ₁₃ | C ₁₄ | C ₁₅ | C ₂₁ | C ₂₂ | C ₂₃ | C ₃₁ | C ₃₂ | C ₃₃ |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| C ₁₁ | 0.132 | 0.149 | 0.089 | 0.160 | 0.191 | 0.206 | 0.235 | 0.123 | 0.166 | 0.218 | 0.182 |
| C ₁₂ | 0.193 | 0.120 | 0.137 | 0.177 | 0.197 | 0.149 | 0.201 | 0.134 | 0.151 | 0.168 | 0.143 |
| C ₁₃ | 0.322 | 0.298 | 0.137 | 0.298 | 0.276 | 0.325 | 0.335 | 0.241 | 0.324 | 0.320 | 0.292 |
| C ₁₄ | 0.264 | 0.252 | 0.152 | 0.191 | 0.200 | 0.291 | 0.276 | 0.184 | 0.269 | 0.306 | 0.284 |
| C ₁₅ | 0.246 | 0.254 | 0.145 | 0.214 | 0.150 | 0.175 | 0.246 | 0.162 | 0.251 | 0.270 | 0.174 |
| C ₂₁ | 0.221 | 0.217 | 0.182 | 0.265 | 0.188 | 0.181 | 0.281 | 0.181 | 0.258 | 0.269 | 0.247 |
| C ₂₂ | 0.257 | 0.277 | 0.121 | 0.262 | 0.244 | 0.234 | 0.219 | 0.242 | 0.264 | 0.282 | 0.275 |
| C ₂₃ | 0.134 | 0.152 | 0.100 | 0.147 | 0.124 | 0.157 | 0.248 | 0.116 | 0.206 | 0.211 | 0.204 |
| C ₃₁ | 0.302 | 0.309 | 0.190 | 0.301 | 0.275 | 0.318 | 0.356 | 0.280 | 0.225 | 0.301 | 0.288 |
| C ₃₂ | 0.344 | 0.304 | 0.224 | 0.340 | 0.316 | 0.343 | 0.383 | 0.302 | 0.330 | 0.268 | 0.314 |
| C ₃₃ | 0.297 | 0.224 | 0.170 | 0.318 | 0.240 | 0.323 | 0.357 | 0.287 | 0.326 | 0.341 | 0.218 |

Table A3. Sum of influences given/received on criteria.

| Criteria | r | s | r+s | r-s |
|-----------------|--------|--------|--------|---------|
| C ₁₁ | 0.7210 | 1.1569 | 1.8779 | -0.4359 |
| C ₁₂ | 0.8241 | 1.0727 | 1.8967 | -0.2486 |
| C ₁₃ | 1.3310 | 0.6605 | 1.9914 | 0.6705 |
| C ₁₄ | 1.0584 | 1.0398 | 2.0982 | 0.0187 |
| C ₁₅ | 1.0090 | 1.0136 | 2.0226 | -0.0047 |
| C ₂₁ | 0.6425 | 0.5719 | 1.2144 | 0.0706 |
| C ₂₂ | 0.6957 | 0.7482 | 1.4439 | -0.0525 |
| C ₂₃ | 0.5207 | 0.5388 | 1.0595 | -0.0181 |
| C ₃₁ | 0.8129 | 0.8815 | 1.6944 | -0.0685 |
| C ₃₂ | 0.9121 | 0.9094 | 1.8216 | 0.0027 |
| C ₃₃ | 0.8855 | 0.8197 | 1.7052 | 0.0658 |

References

- Christensen, B.T.; Ball, L.J. Dimensions of creative evaluation: Distinct design and reasoning strategies for aesthetic, functional and originality judgments. *Des. Stud.* **2016**, *45*, 116–136. [\[CrossRef\]](#)
- Ranjan, B.S.C.; Chakrabarti, A. Development and Validation of a Method for Assessment of Novelty and Requirement Satisfaction in Designing. In *International Conference on Research into Design*; Springer: Singapore, 2017; pp. 589–602.
- Sarkar, P.; Chakrabarti, A. Assessing design creativity. *Des. Stud.* **2011**, *32*, 348–383. [\[CrossRef\]](#)
- Demirkan, H.; Afacan, Y. Assessing creativity in design education: Analysis of creativity factors in the first-year design studio. *Des. Stud.* **2012**, *33*, 262–278. [\[CrossRef\]](#)
- Williams, A.; Askland, H.H.; Boud, D. *Assessing Creativity: Strategies and Tools to Support Teaching and Learning in Architecture and Design*; Final Report; Elsevier: Amsterdam, The Netherlands, 2012.
- Cho, J.Y. An investigation of design studio performance in relation to creativity, spatial ability, and visual cognitive style. *Think. Ski. Creat.* **2017**, *23*, 67–78. [\[CrossRef\]](#)
- Onsman, A. Assessing creativity in a ‘New Generation’ architecture degree. *Think. Ski. Creat.* **2016**, *19*, 210–218. [\[CrossRef\]](#)
- Edward, J.M. Creativity: Who, How, Where. In *Knowledge and the Economy*; Springer: Dordrecht, The Netherlands, 2013; pp. 79–93.
- Sternberg, R.J.; Lubart, T.I. The concept of creativity: Prospects and paradigms. In *Handbook of Creativity*; Cambridge University Press: Cambridge, UK, 1999; pp. 3–15.
- Simonton, D.K. Creativity: Cognitive, personal, developmental, and social aspects. *Am. Psychol.* **2000**, *55*, 151. [\[CrossRef\]](#)
- Plucker, J.A.; Beghetto, R.A.; Dow, G.T. Why isn’t creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educ. Psychol.* **2004**, *39*, 83–96. [\[CrossRef\]](#)
- Glăveanu, V.P. Principles for a cultural psychology of creativity. *Cult. Psychol.* **2010**, *16*, 147–163. [\[CrossRef\]](#)

13. Christiaans, H.; Venselaar, K. Creativity in design engineering and the role of knowledge: Modelling the expert. *Int. J. Technol. Des. Educ.* **2005**, *15*, 217–236. [[CrossRef](#)]
14. Yuan, X.; Lee, J.H. Toward a Computational Approach of Creativity Assessment in Product Design. In *International Conference on Computer-Aided Architectural Design Futures*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 50–62.
15. Tzeng, G.-H.; Shen, K.-Y. *New Concepts and Trends of Hybrid Multiple Criteria Decision Making*; CRC Press/Taylor and Francis Group: Boca Raton, FL, USA, 2017.
16. Hu, S.K.; Tzeng, G.H. A Hybrid Multiple-Attribute Decision-Making Model with Modified PROMETHEE for Identifying Optimal Performance-Improvement Strategies for Sustainable Development of a Better Life. *Soc. Indic. Res.* **2019**, 1–33. [[CrossRef](#)]
17. Amabile, T.M. Motivation and creativity: Effects of motivational orientation on creative writers. *J. Personal. Soc. Psychol.* **1983**, *48*, 393–399. [[CrossRef](#)]
18. Torrance, E.P. Predictive validity of the Torrance tests of creative thinking. *J. Creative Behav.* **1972**, *6*, 236–262. [[CrossRef](#)]
19. Sarkar, P.; Chakrabarti, A. Studying engineering design creativity-developing a common definition and associated measures. In Proceedings of the NSF Workshop on Studying Design Creativity, Aix-en-Provence, France, 10–11 March 2008.
20. Christiaans, H.H. Creativity as a design criterion. *Commun. Res. J.* **2002**, *14*, 41–54. [[CrossRef](#)]
21. Dugosh, K.L.; Paulus, P.B.; Roland, E.J.; Yang, H.C. Cognitive stimulation in brainstorming. *J. Personal. Soc. Psychol.* **2000**, *79*, 722. [[CrossRef](#)]
22. Dugosh, K.L.; Paulus, P.B. Cognitive and social comparison processes in brainstorming. *J. Exp. Soc. Psychol.* **2005**, *41*, 313–320. [[CrossRef](#)]
23. Liikkanen, L.A.; Perttula, M. Inspiring design idea generation: Insights from a memory-search perspective. *J. Eng. Des.* **2010**, *21*, 545–560. [[CrossRef](#)]
24. Jansson, D.G.; Smith, S.M. Design fixation. *Des. Stud.* **1991**, *12*, 3–11. [[CrossRef](#)]
25. Dahl, D.W.; Moreau, P. The influence and value of analogical thinking during new product ideation. *J. Mark. Res.* **2002**, *39*, 47–60. [[CrossRef](#)]
26. Linsey, J.S.; Tseng, I.; Fu, K.; Cagan, J.; Wood, K.L.; Schunn, C. A study of design fixation, its mitigation and perception in engineering design faculty. *J. Mech. Des.* **2010**, *132*, 041003. [[CrossRef](#)]
27. Glăveanu, V.P. The cultural genesis of creativity: An emerging paradigm. *Rev. Psihol. Școlară* **2009**, *2*, 50–63.
28. Norton, D.; Heath, D.; Ventura, D. Finding creativity in an artificial artist. *J. Creat. Behav.* **2013**, *47*, 106–124. [[CrossRef](#)]
29. Zhu, B.W.; Zhang, J.R.; Tzeng, G.H.; Huang, S.L.; Xiong, L. Public Open Space Development for Elderly People by Using the DANP-V Model to Establish Continuous Improvement Strategies towards a Sustainable and Healthy Aging Society. *Sustainability* **2017**, *9*, 420. [[CrossRef](#)]
30. Xiong, L.; Teng, C.L.; Zhu, B.W.; Tzeng, G.H.; Huang, S.L. Using the D-DANP-mV Model to Explore the Continuous System Improvement Strategy for Sustainable Development of Creative Communities. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1309. [[CrossRef](#)] [[PubMed](#)]
31. Qu, G.B.; Zhao, T.Y.; Zhu, B.W.; Tzeng, G.H.; Huang, S.L. Use of a Modified DANP-mV Model to Improve Quality of Life in Rural Residents: The Empirical Case of Xingshisi Village, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 153. [[CrossRef](#)] [[PubMed](#)]
32. Liou, J.; Lu, M.T.; Hu, S.K.; Cheng, C.H.; Chuang, Y.C. A hybrid MCDM model for improving the electronic health record to better serve client needs. *Sustainability* **2017**, *9*, 1819. [[CrossRef](#)]
33. Chuang, Y.C.; Hu, S.K.; Liou, J.J.; Lo, H.W. Building a Decision Dashboard for Improving Green Supply Chain Management. *Int. J. Inf. Technol. Decis. Mak.* **2018**, *17*, 1363–1398. [[CrossRef](#)]
34. Gölcük, İ.; Baykasoğlu, A. An analysis of DEMATEL approaches for criteria interaction handling within ANP. *Expert Syst. Appl.* **2016**, *46*, 346–366. [[CrossRef](#)]
35. Campanella, G.; Ribeiro, R.A. A framework for dynamic multiple-criteria decision making. *Decis. Support Syst.* **2011**, *52*, 52–60. [[CrossRef](#)]
36. Peng, K.H.; Tzeng, G.H. A hybrid dynamic MADM model for problem-improvement in economics and business. *Technol. Econ. Dev. Econ.* **2013**, *19*, 638–660. [[CrossRef](#)]
37. Hall, W.A.; Callery, P. Enhancing the rigor of grounded theory: Incorporating reflexivity and relationality. *Qual. Health Res.* **2001**, *11*, 257–272. [[CrossRef](#)] [[PubMed](#)]

38. Strauss, A.L.; Corbin, J.M. *Shaping a New Health Care System. The Explosion of Chronic Illness as a Catalyst for Change*; Jossey-Bass Publisher: San Francisco, CA, USA, 1988.
39. Strauss, A.L.; Corbin, J.M. *Basis of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*; Sage Publications: Newbury Park, CA, USA, 1998.
40. Corbin, J.M.; Strauss, A. Grounded theory research: Procedures, canons, and evaluative criteria. *Qual. Sociol.* **1990**, *13*, 3–21. [[CrossRef](#)]
41. Glaser, B.; Strauss, A. Grounded theory: The discovery of grounded theory. *Sociol. J. Br. Sociol. Assoc.* **1967**, *12*, 27–49.
42. Liu, F.; Kang, J.A. grounded theory approach to the subjective understanding of urban soundscape in Sheffield. *Cities* **2016**, *50*, 28–39. [[CrossRef](#)]
43. Charmaz, K. Teaching theory construction with initial grounded theory tools: A reflection on lessons and learning. *Qual. Health Res.* **2015**, *25*, 1610–1622. [[CrossRef](#)] [[PubMed](#)]
44. Lovell, J.L. How parents process child health and nutrition information: A grounded theory model. *Appetite* **2016**, *97*, 138–145. [[CrossRef](#)] [[PubMed](#)]
45. Strauss, A.; Corbin, J. *Basic of Qualitation Research: Grounded Theory Procedures & Techniques*; Sage: Thousand Oaks, CA, USA, 1990.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).