BH-ShaDe: A Software Tool that Assists Architecture Students in the Ill-Structured Task of Housing Design

Eva Millán, María-Victoria Belmonte, Manuela Ruiz-Montiel, Juan Gavilanes, and José-Luis Pérez-de-la-Cruz

Abstract—In this paper, we present BH-ShaDe, a new software tool to assist architecture students learning the ill-structured domain/task of housing design. The software tool provides students with automatic or interactively generated floor plan schemas for basic houses. The students can then use the generated schemas as initial seeds to develop complete residential projects. The main goal of our research was to obtain evidence about whether or not such schemas can be useful to architecture students. A first prototype of the tool was evaluated with 78 students, with positive results. However, the students seemed to demand increased user participation, so they could contribute to generating better quality starting points. A second prototype was therefore implemented, allowing a higher degree of interactivity. The second prototype was evaluated with a new group of 50 students. From the two evaluations performed, it can be concluded that both versions of the tool were able to generate useful starting points (either automatically or interactively) that expedited the design process. Additionally, in the second experiment, we found that neither the nature (automatic or interactive) nor the quality of the starting point seems to have any effect on the perceived quality of the final projects.

Index Terms—Computer-aided design, computer-assisted Instruction

1 Introduction

Design is a complex task, that requires a certain set of abilities and skills, inspiration and creativity. In particular, housing design is commonly used as an example of both an ill-structured domain [1] and an ill-structured task/problem [2], [3]. In effect, it presents all of the features that, according to [4] make a problem/task ill-structured: indefinite starting point, indefinite ending point, and unclear strategies for finding solutions. It also presents all the characteristics that, according to [3] can be used to define ill-defined domains: it has multiple and controversial solutions; there is no complete formal domain theory; the associated tasks are also ill-defined; it relies on open-textured concepts and it cannot be divided into smaller independent problems. In the continuum defined in [5] between well-defined and ill-defined problems, housing design belongs to the more challenging category: *class 5*, where multiple solution strategies exists and solution correctness cannot be verified automatically.

For these reasons, housing design is a difficult subject to teach/learn. It is also hard to find software tools that could assist designers beyond the traditional Computer Aided Design (CAD) tools, whose main goal is the creation, modification, or analysis of drawings for a given design.

- *E. Millán, M.V. Belmonte, M. Ruiz-Montiel, and J.L. Pérez-de-la-Cruz are with the Escuela Técnica Superior de Ingeniería Informática, Universidad de Málaga, Málaga, Spain. E-mail: {eva, mavi, mruiz, perez}@lcc.uma.es.*
- *J. Gavilanes is with the Escuela de Arquitectura, Universidad de Málaga, Málaga, Spain. E-mail: jgavilanes@uma.es.*

Manuscript received 24 July 2015; revised 3 Mar. 2016; accepted 23 Mar. 2016. Date of publication 20 Apr. 2016; date of current version 14 Sept. 2016. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TLT.2016.2556667

Building intelligent tutoring systems for ill-structured domains has been a topic that has received attention from researchers in recent years. Some examples for other ill-structured domains (for example medical diagnosis or legal argumentation, both of them *class 5*) are described in [6]. More recently, a multiparadigm intelligent tutoring system for robotic arm training (*class 4*) has been developed and evaluated [7]. However, as far as we know, currently there is no intelligent tutoring system for housing design. The reason might be that developing ITS for this domain is a such a challenging task that other educational technology approaches may be in order.

In the research presented here we have implemented and evaluated two prototypes of a software tool called BH-ShaDe (Basic House Shape Design) whose main goal is to assist architecture students in the task of housing design. BH-ShaDe tries to go beyond traditional CAD tools in helping students design residential projects. To do so, BH-ShaDe generates basic house schemas that can serve as starting points or initial seeds for residential projects. These basic house schemas are used by students as inspiration sources. The concept of *inspiration source* is not new. It refers to the conscious use of different resources or even previous designs, as references for the solution to a problem [8]. Inspiration sources can be very different in nature and play different roles in the design process.

Each prototype has been evaluated with a different cohort of students: the first version of the tool was evaluated with a group of 78 students (first cohort or Group A, in what follows) and the second version with a new group of 50 students (second cohort or Group B). These cohorts belonged to successive academic courses. The main research question to be answered in these two evaluations is:

Research question 1. Can the basic house schemas provided by BH-ShaDe be helpful for students of architecture, in the task of designing residential projects?

The first prototype of BH-ShaDe generated the basic house schemas automatically. In the first evaluation with Group A, the students found that the schemas generated by the tool were useful as starting points for their housing designs. They also valued positively the randomness and variety of solutions provided by the tool, which gave birth to creative projects. However, they also demanded a greater degree of control in the initial solution (i.e., they wanted to participate in its design).

To accommodate these results, we built the second prototype, which could be used in two different working modes: automatic and interactive. Therefore in the second evaluation we were interested in obtaining additional results of the new working mode. To this end, we added two more research questions:

Research question 2. Do students have any preferences for interactive or automatic solutions?

Research question 3. Does the nature (automatic/interactive) or quality of the initial seed have any impact on the perceived quality of the final residential projects presented by the students?

In this second evaluation we also obtained good results for our first research question. The evidence suggests that starting points provided by the tool were useful. For the additional research goals established in this second experiment, results indicated that the benefits of increased user control (demanded by the first cohort of the students) seem to be somewhat unclear. The results of the experiments suggest that the students did not show a clear preference for either of the two working modes and that they were able to produce good quality projects using the initial solutions provided by the tool, independently of its nature and quality.

2 Literature Review

In this section we review the background and related work that we have used as the theoretical basis of our approach. We begin by analyzing former uses of computers in architectural design education. As our software tool relies on the formalism of shape grammars, we briefly outline some computational techniques that have been used in design, and then present shape grammars and some interesting applications in educational processes. Finally, as we use basic house schemas generated by the tool as inspiration sources for the students, we present some other approaches which also use inspiration sources in different educational contexts related to design processes.

2.1 Software Tools in Architectural Education

First, we begin by analyzing how computers and software tools have been used in architectural education. In this vein, Andia presented an interesting study about the influence of computers in both professional practice and architecture teaching [9]. According to this study, architects have used computers mainly as a support for functions that have been common practice for the last 150 years. In contrast, in architecture schools, computers have been introduced in more creative ways: experimental laboratories for design, as an aid to imagination, to assist teaching and learning, and also to introduce virtual reality into architectural education. The importance of using computers in an innovative way in architectural practice has been advocated by some authors [10]. One of the options to do this is the so-called *algorithmic architecture* or *modeling generative architecture* [11], which combines architectural design and artificial intelligence to develop algorithms that simplify or automate design and planning tasks. Next we present some examples of successful uses of computers in architectural education, that go beyond just using software tools for digital drawing.

We can find few but interesting examples of software tools that, like BH-ShaDe, have been developed for architectural design education and evaluated with real students. An example is DYNAMO [12], which is a web-based design assistant to support architectural design education. DYNAMO helps students by presenting an on-line collection of design cases, that provide ideas for their own projects. The tool was evaluated with 48 students. The researchers concluded that students found the tool engaging, however they did not exploit the opportunities of active participation (i.e., added new cases or commented on existing ones). Another example is SketSha [13], that provides support to free hand sketching in locations far from each other, to allow for collaborative design. It was used by 38 students in different locations (Belgium and France) and, according to the authors, the experience was a success in terms of the quality of the projects and the students' satisfaction.

A quite recent and innovative use of computers in design education is based upon Augmented Reality (AR) systems. For example, in [14], AR systems were used both to explain the relevant domain knowledge of creative design and as a test-bed so that secondary-school students could build their own AR scenes. The results of the study they carried out (with an experimental group of 19 students and a control group of 18 students) suggested that in the experimental group, this learning scheme improved the students' attention and motivation and enhanced the creativity of their final designs, with respect to the control group.

Another interesting and recent proposal can be found in [15]. They use a prompt-based annotation approach to support mobile learning activities in architectural design. An experiment was conducted, involving 28 students in the experimental group and 21 in the control group. Students in the control group learned with a traditional approach that involved taking photos and paper and pencil annotations, while students of the experimental group learned with the mobile learning approach with the prompt-based annotation. The results showed that the approach promoted students self-efficacy, increased the cognitive load during the learning activity and improved the learning achievement.

2.2 Shape Grammars in Architectural Education

From the standpoint of the application of computational techniques in architectural education, there has been increasing interest in developing "intelligent" (i.e., adaptive) software tools for learning. However, to the best of our knowledge there are no intelligent tutoring systems (ITSs) or intelligent learning environments (ILEs) for architectural design. As explained, the difficulty of implementing ITSs or

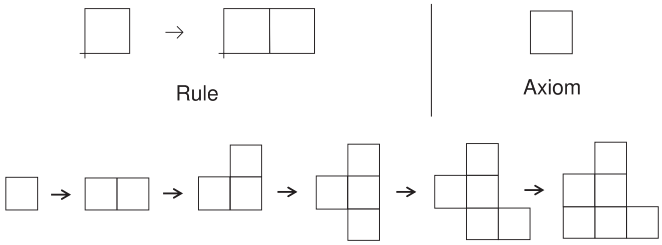

Fig. 1. Shape grammar that adds squares, with one possible derivation.

ILEs for this domain might arise from the fact that it is ill-structured. The absence of a well-defined sequence of steps to produce good designs makes it difficult to implement this kind of educational software tools. BH-ShaDe is not an ITS either. However, it uses artificial intelligence techniques to automatically generate the basic house schemas.

Concerning computational techniques, several approaches have been used to stimulate the creativity and exploration in design: genetic algorithms [16], case based reasoning [17], analogy [18] or shape grammars ([19], [20]). Shape grammars have been widely used in design, and, specifically, in architectural design [21]. In our proposal we have used shape grammars, which we briefly describe next.

A shape grammar is a formal language that represents visual thinking. To this end, there is an initial shape (usually called *the axiom*), and a set of design rules or transformations that can be applied to different shapes. Fig. 1 represents an example of an axiom, a rule, and five successive applications of the rule, starting from the original axiom.

These transformations are applied iteratively in random order and localizations in the design. Therefore, these techniques are suitable for applying randomness to generate a variety of forms, which might be an interesting approach to stimulate visual thinking. In this way, in the work presented here, shape grammars are used as a generating device that facilitates the production of a large amount of starting points that can stimulate and expedite the design of schemas for floor plans of single-family dwellings. Reinforcement learning techniques have been used to control the quality of the solutions generated.

Regarding the use of shape grammars for architectural design education, some of the more significant examples are the design projects developed by the students of Terry Knight at the Massachusetts Institute of Technology and the University of California, Los Angeles. For example, Randy Brown (UCLA, 1980) used a simple 3D shape grammar to design a museum of cultural history.

Another example is the Master's thesis project by Knight's student McGill [22], which shows the role of computers as facilitators for learning about shape grammars and their use in the architectural design process. For this study, a software prototype, Shaper2D, was created. To test this prototype, three studies with undergraduate architecture students at the MIT Department of Architecture were conducted. In these studies, Shaper2D was used to generate the site layout and massing for family townhouses. The results showed that although Shaper2D could be used as a design tool, it was only useful when the designer did not have a fixed idea, and that using the tool to look for a pre-conceived design could cause frustration.

In [23], shape grammars have also been used in an educational software tool. In this case, the tool generates 3D-megastructures. The tool was used in an experiment with 62 architecture students. The results show that the students were able to generate good solutions by exploring the randomness provided by the tool, but also liked to keep a certain degree of control over their designs. Our approach is similar to this one, in the sense that the inspiration sources are generated by the software tool, and that to do so, the tool relies on the use of shape grammars. However our software tool uses shape grammars to generate basic house schemas, instead of 3D-megastructures.

2.3 Starting Points in Architectural Design Education

Finally and concerning the use of starting points for teaching design, there are several studies in the related literature. For example, Iordanova performed a study with 10 architecture students that used a library of digital models as a source of inspiration or referents during their work on a design task [24]. In the evaluation, the results showed that 38 of the 50 design projects were linked to the use of the referents. Furthermore, the authors reported a low degree of imitation/copy of the original in favor of a higher degree of the use of the *know-how* embodied in the referents library.

Sketches have also been used as visual stimuli [25]. Three groups of 12 architecture/industrial design students solved four design tasks under different conditions; no visual stimuli, (group 1); rich, diverse visual stimuli (group 2) and modest visual stimuli (group 3). The results showed that, when students are required to solve ill-structured design problems, the presence and nature of visual stimuli does have an impact on the quality of the solutions.

In another study, the researchers provided 17 experienced designers and 22 architecture students with a rich collection of visual displays [26]. The subjects in the experimental group were explicitly asked to use these displays as potential analogues for their designs, while the subjects in the control group were not. The results of the experiment showed that subjects (of any level of expertise) who are provided with visual displays use them to enhance their design problem ability, and that the use of visual analogies produced better design results in both groups.

In this sense, we have also tried to use the basic house schemas as visual stimuli for the students to solve an ill-defined task. These external stimuli are expected to act as triggers for the generation of ideas and provoke the creative leap that occurs between the problem and the solution [27]. In [28] it is showed that when ill-structured design problems are approached through exploration of design alternatives, creativity is fostered and learning enhanced.

With respect to the ideal number of initial solutions to be considered as a source of inspiration by designers, there seems to be different views. Some studies suggest that working with more alternative solutions contributed to better performance [29], while other studies advocate that "less is more (original)" [30]. Our approach in this regard is similar to the one presented in [31]: we produce many different alternative solutions (as many as required, as the tool has generative capabilities) but propose to keep the number controlled by a posterior and conscious evaluation and assessment of such alternatives.

After this literature review, we can conclude that recent trends point to the use of computer tools as part of the creative process. Some authors propose the use new computational approaches, like *modeling generative architecture*. The computational power of digital media can therefore be incorporated into architectural design.

In this research context, BH-ShaDe is a software tool that has been specifically implemented with the purpose of exploring the possibilities and usefulness of tools that are able to (automatically or interactively) generate basic house schemas to provide support to residential design projects. To do this, it makes use of artificial intelligence techniques and shape grammars. Such initial seeds can be used as inspiration sources or referents, and students can use them as starting points for their residential projects.

3 Methodology

In this section we describe the methodology we have used.

3.1 Participants and Environment

The experiments were carried out in two successive academic courses. Each year the tool was tested with the students of the subject *Architectural Projects VII*, which is taught in the seventh semester of the five-year Architecture Degree of the University of Málaga. The first cohort (first academic year) had 78 students and the second cohort (second academic year) had 50. All the participants had experience in designing residential projects (acquired in previous subjects). Each academic year students are divided into three teaching groups. Each of these groups is taught by a different tutor. One of them is a member of our research team (and participated in the development of the software tool), while the other two did not have previous experience with these kinds of tools.

The learning activity proposed to the students was carried out during three sessions of two hours each. That is, it involved six hours of student's work in the classroom. It also included two assignments that the students solved as homework (outside the classroom). It was identical for both cohorts: designing single-family dwelling houses and three types of groupings: apartments blocks, row-houses and galleries. To this end, students had to use the schemas generated by our software tool as starting points to develop complete residential projects. Each cohort worked with a different prototype of the tool, the main difference being that the second prototype could also be used in interactive mode to generate basic house schemas. Both prototypes were able to automatically generate any given number of basic house schemas.

The learning activity was compulsory for all students. It was evaluated by their tutors, and it was a component of the final mark. It was not possible to separate students into control and experimental groups, due to academic regulations at our university.

3.2 Software Tool: BH-ShaDe

In this section we briefly describe the software tool we have developed for this study. BH-Shade generates and proposes housing units schemes that can serve as starting points in students' exercises and projects. The tool has been implemented on Trimble SketchUp (`http://www.sketchup.com/`), and is based on the ideas of reinforcement learning and shape grammars. A complete description of the tool can be found in [32], and a user's guide is available on the web [33].

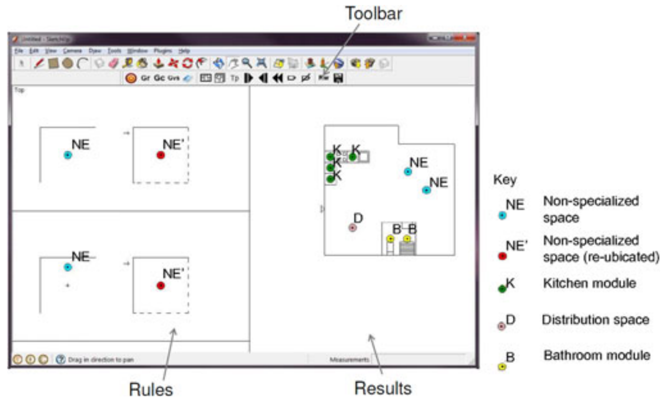

Fig. 2. Screenshot of BH-ShaDe interface.

BH-ShaDe integrates a fixed shape grammar and generates its output according to it. This grammar implements a housing program developed by an architectural studio [34] for the regional government of Andalucía (Spain), that specifies the criteria that a basic house must meet, depending on the number of inhabitants.

The output of BH-ShaDe is a number of two-dimensional floor plan distribution schemas (in the following, *schemas* or *cells* or *seeds*) of basic, two-person housing units. All the schemas produced are distributed over one floor and its total area is restricted to $46\ m^2$. In the housing proposal, several kinds of spaces are considered: (1) *specialized spaces* (which need specific installations), (2) *non-specialized spaces* (do not need specific installations, and their use is determined by its inhabitants: dining-room, living-room, bedroom) and (3) *complementary spaces*, such as distribution hall, that allows circulation between spaces. In Fig. 2 the BH-ShaDe interface with a generated schema is shown.

However, the solutions generated by shape grammars are usually not feasible. For example, see the first solution in Fig. 3, which clearly shows that further control mechanisms are needed. To this end, we have used reinforcement learning techniques to produce good quality solutions (like the second solution in Fig. 3. A complete description of the shape grammar and the reinforcement learning policy used can be found in [35].

As explained, we have implemented and evaluated two different prototypes of the tool. The first prototype only allows automatic generation of the basic house schemas, while the second also allows interactive solutions (as demanded by students after the first experiment). Let us explain in more detail the two different working modes:

- Automatic mode: the student specifies the desired number of schemas and BH-ShaDe will generate them.
- Interactive mode: the student can interact with the tool by supervising the outcomes of the different steps (contour generation, placing of the distribution

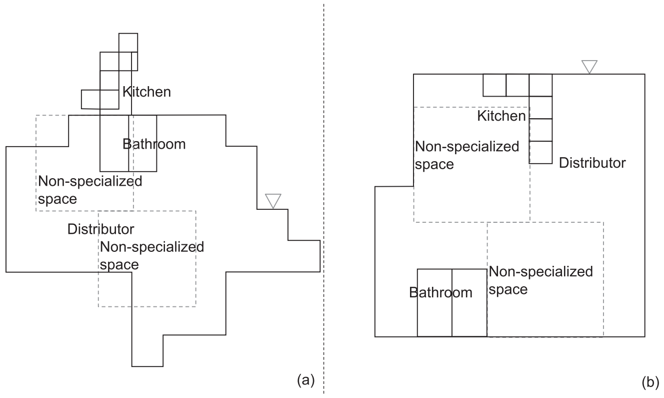

Fig. 3. Schemas without (a) and with (b) reinforcement learning.

hall, or kitchen generation). If the student is not satisfied with the result, he/she can repeat the step or go back in the sequence.

3.3 Learning Activity

In this section we describe the learning activity given to the students, which was identical in the two groups. The only difference is that Group A used automatically generated schemas, while Group B used a mixture of interactive and automatic ones. The complete activity involved three two-hours sessions. The time lapse between sessions was a week.

3.3.1 First Session

The students received a lesson on shape grammars (and their use in the design process) and about the housing program described in [34]. Then they had the opportunity to individually use BH-ShaDe in some simple interactive exercises, involving the use of the two simple grammars presented above. The goal of this session was that the students could understand and learn about shape grammars and also about how to use BH-ShaDe.

The students were then divided into groups. In the first cohort, students were divided into nine groups of 8-9 students each. Each group then used the tool to automatically generate 81 basic house schemas. In the second cohort the students were divided into nine groups of 5-6 students each. Each group generated 72 schemas automatically and 18 interactively, resulting in a total of 90 schemas.

Then the groups of both cohorts were given an identical learning activity, which consisted in analyzing and discussing the perceived quality of the initial seeds. This learning activity was carried out by the students as an assignment (out of the class). The number of schemas was increased for the second cohort because the tutors considered that it had been a very good exercise in analysis and reflection for students of the first cohort.

This reflection process concluded with a classification of each schema into the one of the following categories: A (optimal), B (adequate), C (some modifications needed), D (problematic) and E (absurd). The students had to agree on their criteria for the classification of the schemas, as part of their self-reflection process (first individually, and then as a group) to learn what a "good quality" schema is.

Then, the students had to select the initial seeds to be used in their residential projects. No instructions were given to them, neither about the number nor about the characteristics (generation mode or quality) of the initial seeds to be chosen. In this way, the students were allowed to freely explore the complete range of solutions generated by (or developed with the help of) the tool. The students had to select schemas for four different residential projects: single-family houses and three different types of groupings: apartment blocks, row houses and galleries.

3.3.2 Second Session

The students presented their work to the instructors and to their colleagues. In their presentation they had to explain the criteria used to evaluate the schemas (from A-E), and also to show the first drafts of their residential projects (including the initial seeds selected). During the session they received feedback and questions from their tutors, which they used to finish their residential projects (out of the class).

3.3.3 Third Session

In the last session, the students had to present their final designs (a total of 36 residential projects per cohort). For example, Fig. 4 shows a single-family dwelling (based in a A schema), while Fig. 5 displays an apartment block (based on two D schemas). In these figures we can see:

- Figs. 4a, 5a and 5b: the initial schemas, the modifications made to them and the final floor plan;
- Fig. 5c: the grouping pattern, and
- Figs. 4b, 4c, 4d and 5d: some views of the final residential projects.

These examples of the student's projects have been chosen to illustrate something that will be shown later (in the discussion of research question 3): for single-family houses, the students showed a preference for the best starting points (with few exceptions), while for groupings they exploited the complete range of initial seeds by choosing some of the worst ones (and making the necessary adjustments).

Right after each presentation, the tutors discussed and evaluated the projects. Finally, after completing the task and being evaluated for it, the students completed a small questionnaire (about the task and the tool).

Recall that housing design is a *class 5* task/domain. Therefore this learning activity does not include an automated analysis of students solution. However, it promotes the use of suitable teaching techniques for ill-structured domains, as described in [3]: a) *case studies* (students learn from the discussions and feedback provided by their instructors for all projects); b) *weak theory scaffolding* (through the housing program presented in [34]; c) *expert review* (by means of the discussions and feedback from instructors) and some sort of d) *peer review collaboration* (students reviewed the quality of the schemas, both of those generated automatically by the tool, but also of those generated interactively by their peers). We have also introduced computers in the learning activity in an innovative way, by means of the generative capabilities and interactivity possibilities of the software tool.

3.4 Study Design

As explained in the introduction, the main research question for our study is:

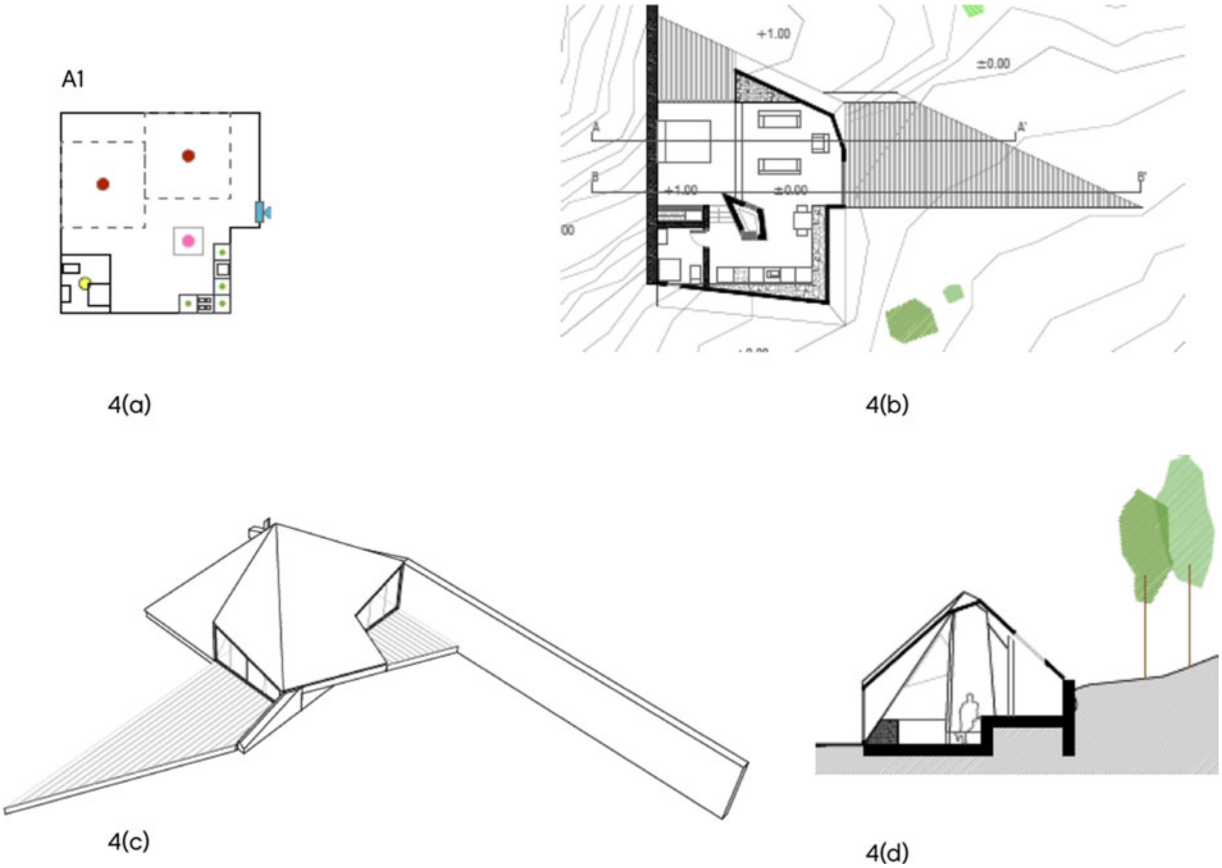

Fig. 4. A single-family dwelling.

Fig. 5. An apartment block.

Research question 1. Can the basic house schemas provided by BH-ShaDe be helpful for students of architecture, in the task of designing residential projects?

After the first evaluation we used the results to develop a second prototype of the tool, and two more research questions were added.

Research question 2. Do students have any preferences for interactive or automatic solutions?

Research question 3. Does the nature (automatic/interactive) or quality of the initial seed have any impact on the perceived quality of the final residential projects presented by the students?

Research questions 1 and 2 are qualitative, while research question 3 is quantitative. The variables we have defined for research 3 are: I (degree of interactivity of the initial seed); q (quality of the initial seed) and Q (quality of the residential project). Next we describe the measure instruments used in the study: a questionnaire for students, a questionnaire for the instructors participating in the experiment, and the student's projects.

The questionnaire for students was designed to obtain their opinion about several aspects: the software tool itself, the quality of the schemas proposed by the tool, the learning task, and their preferences for the interactive or automatic modes. To this end, it included types of questions, namely:

- Twelve Likert items, relative to four different topics: software tool, basic house schemas (and their usefulness as starting points), global evaluation, and the practice. The complete questionnaire and its results are shown in Section 4.1.

  The students had to evaluate their degree of agreement with the twelve sentences from 1 (strongly disagree) to 6 (strongly agree). An even number of possible answers was chosen for the Likert items, to avoid the central tendency bias [36].

  In the second experiment, three more Likert items were added to the questionnaire, intended to obtain evidence about the students' preferences (P) about automatic/interactive initial seeds (I).
- Two free-text items, where students could identify the strong and weak points of the software and make any comments they wished.
- A multiple choice item, so the students could explicitly express their preference for either of the two different working modes (automatic versus interactive). This multiple choice question was only included in the questionnaire for the students of Group B (who used the second prototype of the tool).

The questionnaire for instructors is composed of four questions, aimed to obtain their opinion about the software tool and the learning activity. The two tutors were asked to fill the questionnaire together and provide a single answer to each question. The complete questionnaire and its results are shown in Section 4.3.

Finally, student's final projects were used also as a measure instrument to address the research questions.

3.5 Methodology for the Analysis of the Results

In this section we describe how the results of each measure instrument have been analyzed.

The students' answers to Likert items in the questionnaire were evaluated using the recommended methodology [37] to analyze Likert items, that is, mode, median, interquartile range and nominal levels of disagreement (degree of disagreement of 1, 2 and 3) versus agreement (degree of agreement of 4, 5 and 6).

Free-text items were processed according to the *constant comparative method* [38], a methodology based on *grounded theory* [39]. In the first step, each student's response was decomposed into the ideas it expresses (*answers*). Therefore there are usually more *answers* than students, because each student's response usually expresses more than one idea. Then the *answers* were divided into *categories*. In the *phenomenological reduction phase*, the *categories* were grouped by subject (*themes*). Finally, in the *triangulation phase*, examples of *supporting quotes* were provided. The main advantage of using this methodology is that the ideas expressed in students' answers emerge from the analysis of the sentences, and are not pre-conceived by the researchers. Therefore the free text items are not intended to measure any predefined variable, but to collect information that arises from the students answers.

Next we describe the methodology for the evaluation and analysis of the students' projects. As explained before, in the first evaluation of the software tool with Group A, the students seemed to demand a greater degree of interactivity (that allowed them to generate more appropriate initial seeds). To accommodate this, a second prototype with an interactive mode was implemented. To evaluate the usefulness of this new working mode, we carried out a study with the projects presented by the second cohort. The main goal of this study was to determine whether or not the greater degree of interactivity in the second prototype of the tool had had any effect on the perceived quality of the final residential projects. Another goal was to analyze the possible influence of the quality of the initial seeds in the quality of the projects (Q). To this end, we analyzed each of the 36 projects of the second cohort, according to three dimensions: degree of interactivity of starting points (I), quality of starting points (q), and quality of the final design (Q). Here we describe each dimension in more detail.

- Degree of interactivity of the starting point (I). As explained, each group had absolute freedom to select from their 90 schemas those to be used as starting points. Therefore, for each project we can define a *degree of interactivity of the starting points* I, that we define as the ratio between interactive starting points and total number of starting points used in that particular project.
- Quality of the starting point (q). Each starting point was labelled from A (optimal) to E (absurd). For each project, let q_w be the worst value of these labels and q_b the best one.
- Quality of the design (Q). In order to assess the overall quality of the designs submitted by the students, we have used the five categories defined in [40] to score the quality of the designs: creativity (Q_c), aesthetics (Q_a), ergonomics (Q_e), technical aspects (Q_t) and business aspects (Q_b). The global quality of each project, Q is computed as the average of these five variables. Each project was assessed independently

TABLE 1
Results of the Students' Questionnaire

	Group A			Group B			Group A	Group B
	Med.	Mod.	(Q_1, Q_3)	Med.	Mod.	(Q_1, Q_3)	% Agree	% Agree
About the software tool								
1. I quickly learned how to use the tool	5	5	(5,6)	5	6	(5,6)	91.03	92.86
2. It was easy for me to use the tool	5	6	(5,6)	5	6	(5,6)	97.44	100
3. The user interface is intuitive	4	4	(4,5)	5	5	(4,5)	79.49	88.10
4. The tool worked sufficiently quickly	4	4	(3,5)	5	5	(3,5)	71.79	73.81
About the schemas proposed by the tool...								
5. The schemas can provide good starting points	4	5	(4,5)	5	5	(4,5)	75.64	83.33
6. The schemas were interesting	4	4	(3,5)	4	5	(3,5)	62.82	71.43
Global Evaluation								
7. Without the tool, this task would have been more difficult	3	5	(2,5)	4	4	(3,5)	48.72	59.52
8. I would like to know more about this kind of tools	4	4	(3,4)	5	5	(3,5)	61.54	83.33
9. I would like to be able to define my own shape grammars	3	3	(3,4)	5	5	(3,5)	48.72	71.43
About this task								
10. All in one, it was interesting	4	5	(4,5)	5	5	(4,5)	80.77	92.86
11. I think that the methodology used was suitable	4	4	(4,5)	4	5	(3,5)	75.64	73.81
12. It was rewarding to work in groups	5	5	(4,6)	5	5	(5,6)	79.49	90.48
Interactive mode (only for students of the second cohort)								
13. I liked to be able to combine randomness and control in design	-	-	-	4	4	(3,4)	-	75.19
14. In the automatic mode, randomness can trigger the creative leap	-	-	-	5	5	(4,6)	-	90.48
15. It was useful to have certain degree of control in the interactive mode	-	-	-	5	5	(4,5)	-	92.86

by the two tutors involved in the experiments. Both of them have extensive experience (more than 20 years) in teaching and professional practice. The evaluation of the quality of each project was blind with respect to the quality or generation mode of the initial seeds. The inter-rater agreement was analyzed. The average size of the confidence interval for the difference of scores (across the five different dimensions evaluated) was 0.79, which shows a reasonable degree of agreement. Finally, the final score for each project was computed as the average of the two individual scores.

To further clarify these dimensions, let us show some examples. The single family house shown in Fig. 4 was developed from an automatically generated cell (so $I = 0$) classified as A (therefore $q_w = q_b =$ A). The project was rated by the tutors with $Q_a = 4$ (aesthetics), $Q_t = 3$ (technical aspects), $Q_c = 3.5$ (creativity) and $Q_e = Q_b = 5$ (ergonomic and business). The quality of the project is then the average of such measures, that is, $Q = 4.1$. For the apartment block shown in Fig. 5, students used two automatic schemas that they had classified as D ($I = 0$, $q_w = q_b =$ D). This project also received good scores: $Q_c = Q_a = Q_t = 3.5$ (creativity, aesthetics and technical aspects), and $Q_c = Q_c = 4.5$ (ergonomic and business). Therefore the quality of this project is $Q = 3.9$.

The results of the evaluation of all projects are shown in Table 4. The examples used in the previous paragraph correspond to projects 13 and 14 in such table.

4 Results

In this section we present the results obtained in the two experiments performed. First we begin by reporting the students' answers to the questionnaire (Groups A and B), then the results of the analysis of the residential projects (Group B), and finally the tutors' answers to the questionnaire developed for them.

4.1 Results of Students' Questionnaire

A total number of 78 students (first cohort, Group A) and 42 students (second cohort, Group B) completed the questionnaire. Recall that it was composed by three different types of questions: twelve Likert items and two free text-items (both cohorts). An additional three Likert items and a multiple choice question were also presented to the second cohort. We present the results of each type of question separately.

The first twelve Likert items were answered by both cohorts (i.e., by a total number of 120 students), while the three Likert items specifically for the interactive mode were answered by forty-two students of the second cohort. The students had to evaluate their degree of agreement with each sentence, from 1 (strongly disagree) to 6 (strongly agree). Table 1 summarizes mode, median, inter-quartile range and nominal levels of agreement (degrees of agreement of 4, 5 and 6) for each of the fifteen items.

With respect to the two free-text items, there were 135 (Group A) and 86 (Group B) different *answers* for the positive aspects, and 102 (Group A) and 59 (Group B) different *answers* for the aspects to be improved. Comparatively, the 42 students that completed the questionnaire in the second cohort produced a greater number of different *answers* (both for positive aspects and possible improvements of the tool) than their 78 colleagues in the first cohort.

Two researchers independently assigned *answers* to *categories* (13 in the case of the positive *answers*, and seventeen for the negative ones). The inter-rater agreement between the two researchers was computed using the iota ι statistic [41], an extension of the kappa measure for the case of multivariate data and multiple judges. A ι value of 1 indicates perfect agreement. In our case, we obtained ι values of 0.707 and 0.69 (positive aspects, groups A and B) and 0.75 and 0.898 (aspects to be improved, groups A and B) which indicate a reasonable initial degree of agreement. Then, a negotiation process between the two researchers was completed, to finally assign *answers* to *categories*. The results of applying the constant comparative method to our two free text items are shown in Tables 2 and 3.

Finally, the questionnaire for Group B included a multiple choice question that allowed the 42 students to explicitly state their preferences for the interactive or automatic

TABLE 2
Categories, Themes, and Supporting Quotes for Positive Aspects

	CATEGORIES					THEMES	EXAMPLES OF SUPPORTING QUOTES
	Name	# answers (A vs B)		% answers (A vs B)			
135 different answers (group A) vs 86 different answers (group B)	Diversity	26	16	19.26	18.61	ASPECTS RELATIVE TO QUALITY OF SOLUTIONS (25.18% vs 33.72%)	*"The tool provides numerous alternatives"* *"The tool generated valid dwellings that needed little modification"* *"The outer shape of schemas allows interesting grouping"* *"Automatic generation of multiple schemas with suggestive distributions"*
	Validity	6	4	4.44	4.65		
	Versatility for groupings	2	4	1.48	4.65		
	Suggestive solutions	0	5	0	5.81		
	Usability	3	3	2.22	3.49	ASPECTS RELATIVE TO THE SOFTWARE TOOL (4.44% vs 9.3%)	*"The tool was very easy to use"* *"The tool plays a decisive role in the design process"*
	Using software tools as part of the creative process	3	5	2.22	5.81		
	Possibility of working in groups	10	1	7.41	1.16	TEAMWORK (14.82% vs 12.79%)	*"Working in groups"* *"Having so many alternative initial solutions requires a great deal of analysis and reflection, which requires critical thinking"*
	Processes of selection/reflection carried out in the working groups	10	10	7.41	11.63		
	Randomness	9	3	6.67	3.49	SUPPORT FOR CREATIVITY (55.56% vs 44.19%)	*"Randomness strengthens the design process"* *"Learning that we can take advantage from less-than-perfect schemas"* *"The range of starting points and solutions provided by the tool"* *"Getting ideas for your project that go beyond your preconceived solutions"* *"The tool automatically generates the solutions, helping the designer in this time-consuming task"*
	Happy accidents/bugs	3	4	2.22	4.65		
	The tool provides good starting points	41	19	30.37	22.10		
	Overcoming preconceived solutions	12	7	8.89	8.14		
	Helps designers to be more efficient	10	5	7.41	5.81		

working mode for the generation of the initial seeds. The students could select from three choices: a) I prefer the automatic version, b) I prefer the interactive version or c) I think that each version has its own functionality. The percentage of students that selected each option was 29, 38 and 33 percent, respectively.

4.2 Results of the Analysis and Evaluation of Students Projects

As explained, in the experiment carried out with the second cohort we tried to obtain evidence of the possible influence of the nature and quality of the initial seeds in the perceived quality of the students' projects.

Table 4 shows some results based on Group B students' projects. The first column identifies the project and the second column its type (SF: single-family house; AB: apartment block; RH: row houses; GA: gallery). The third column shows the degree of interactivity of the starting points, I, while the fourth and the fifth columns show the labels for, respectively, the worst (q_w) and best (q_b) schemas used in the project. The rest of the columns show the grading of the projects according to the five selected criteria ($Q_c, \ldots, Q_b$) and its average Q.

As explained, the students had a total of 810 schemas to choose from. From them, 88 were selected by the students as starting points for some of their 36 projects. The percentage of schemas that received each quality rating (q) is shown in Fig. 6.

Concerning the interactivity of the selected starting points, 42 selected schemas were generated automatically and 46 were created interactively. Of the 36 residential projects developed by the students, 14 used automated schemas, 11 used interactive schemas, and 11 projects combined both types.

4.3 Results of Tutors Questionnaire

In this section we present a brief summary of the tutors opinions.

Question 1. What did you like most about this activity and tool?

"The first added value of this activity is simply the possibility of using the tool and knowing the concept of shape grammar. We consider that it is very educational, especially for a generation so accustomed to using software. It is reasonable to think that some students will continue to use them in the future. The fact that this task is interdisciplinary is also very positive."

Question 2. What possible improvements could be made?

"In relation to the tool and its use in this particular activity, we think that it would be desirable to extend the number of variables to be taken into account in the shape grammar, so it would be able to generate solutions that are formally more extreme and less predictable."

Question 3. After seeing the final results, how would you rate the overall work developed by your students? Do you think the results would have been different if students had had to develop their projects from scratch?

"Overall, the students have done a great job, especially taking into account the limited time they had. It is reasonable to assume that even without the starting point provided by the tool, some of the projects presented would have been very similar. We think that some students were looking for their pre-conceived solutions among those provided by the computer tool. However, some others were more open to exploring the different solutions, and we think that the most interesting projects have emerged from accidental elements, like the small annex spaces or errors. Initially, they seemed not to have any practical use, but finally

TABLE 3
Categories, Themes, and Supporting Quotes for Aspects to be Improved

	CATEGORIES					THEMES	EXAMPLES OF SUPPORTING QUOTES
	Name	# answers (A vs B)		% answers (A vs B)			
102 different answers (group A) vs 59 different answers (group B)	Editing capabilities for the software tool (edition)	9	11	8.82	18.65	ASPECTS RELATIVE TO THE SOFTWARE TOOL (12.74% vs 28.82%)	*"The tool should allow the user to: group the dwellings/move the spaces/generate 3D views"* *"A more intuitive user interface"* *"More flexibility in the order used to generate spaces in the interactive mode"*
	Usability	4	-	3.92	-		
	No fixed order for generating spaces in the interactive mode	-	6	-	10.17		
	Shape grammar should include additional criteria	20	2	19.61	3.39	ASPECTS RELATIVE TO SHAPE GAMMARS (21.57% vs 6.78%)	*"The bath door should not be aligned with the front door"* *"I would like to be able to modify rules to generate different types of schemas"*
	Possibility to include user-defined shape grammars	2	2	1.96	3.39		
	Overlapping of non-specialized spaces	11	8	10.78	13.56	ASPECTS RELATIVE TO DISTRIBUTION OF SPACES (48.04% vs 42.37%)	*"Spaces should not overlap, especially when there is free space"* *"Kitchen furniture should be accessible"* *"Avoid residual spaces"* *"The front door does not necessarily need to be near the kitchen"* *"Wet zones should be as near as possible"* *"Better distribution of spaces"*
	Poor distribution of kitchen furniture	3	1	2.94	1.69		
	Residual spaces	10	6	9.81	10.17		
	Better location of the front door	12	8	11.76	13.56		
	Better placement of wet zones	10	2	9.81	3.39		
	Better distribution of spaces	3	-	2.94	-		
	Little variety of solutions	7	6	6.86	10.17	ASPECTS RELATIVE TO SOLUTIONS (17.65% vs 22.03%)	*"The tool should generate a greater variety of solutions"* *"Other parameters should be considered (environmental, social aspects, etc.)"* *"The randomness of the tool should be controlled somehow"* *"E-types should not be shown to the user"* *"Other sizes for the basic module should be allowed"* *"Areas other than $46m^2$ should be considered"*
	Additional criteria should be considered (not only architectural)	2	-	1.97	-		
	Excess of randomness	9	-	8.82	-		
	Absurd solutions should not be shown to the user	-	4	-	6.78		
	No $1m^2$ module	-	2	-	3.39		
	More sizes for dwellings	-	1	-	1.69		

they have served to encourage the clustering of the dwellings, providing support so students could freely use their imagination. We do believe that the use of the tool has accelerated these kind of discoveries and expedited the designs."

Question 4. Please make any comments that you wish.

"About the software tool, it seems clear to us that it is still in the first stages, especially if it were to be used in a professional environment. But in our opinion its use in educative settings is very appropriate. In fact, the tool has been useful at least in providing experiences in two important factors of the design process: the difficulty of making choices and the randomness provided by happy accidents, both of them topics of special interest in our subject matter. Having so many floor plans automatically generated by the tool, so they could be discussed and selected by the groups of students, has been an excellent exercise in analysis and reflection, which are usually easier to carry out in other people's work than in their own designs. At the same time, the program generates such a wide variety of schemas that accidents occurred randomly, creating proposals that at first sight could be considered as undesirable forms, but in the end were used to generate the most interesting projects. Students can therefore learn (in a practical way) that their preconceived ideas are not always the more appropriate solutions."

5 Discussion

In this section we discuss the results we have presented in the preceding section, in terms of the research questions formulated.

Research question 1. Can the basic house schemas provided by BH-Shade be helpful for students of architecture, in the task of designing residential projects?

First of all, the use of the software tool, at least, did not appear to be a hindrance for the development of the design task. In fact, with respect to the usability of the software tool, all related Likert items (1, 2, 3 and 4) show medians and modes of at least 4 and degrees of agreement higher than 71.79 percent across the two cohorts of students.

When asked about the schemas generated by the tool, most students considered that they can provide good starting points (75.64 percent in the first cohort, and 83.33 percent in the second one), and that the schemas were interesting (62.82 and 71.43 percent). It also seems that adding interactivity to the tool improved its usefulness: 59.52 percent of the students of the second cohort considered that "without the tool, the practice would have been more difficult", while in the first cohort 48.72 percent seemed to agree with this sentence.

All in all, the students seemed to enjoy this task: they considered it interesting (80.77 and 92.86 percent) and they

TABLE 4
Analysis and Evaluation of the Projects (Group B)

Project	Type	I	q_w	q_b	Q_c	Q_a	Q_e	Q_t	Q_b	Q
1	SF	1.00	C	C	3.5	3.0	3.0	3.5	3.0	3.2
2	AB	0.67	C	C	3.5	4.5	4.5	3.5	4.0	4.0
3	RH	0.75	C	A	2.5	2.5	2.5	2.0	2.5	2.4
4	GA	0.67	C	C	5.0	4.5	5.0	4.5	5.0	4.8
5	SF	1.00	B	B	4.5	3.5	4.5	4.0	5.0	4.3
6	AB	1.00	B	B	2.5	3.5	3.5	3.0	3.5	3.3
7	RH	1.00	C	B	4.0	4.0	4.5	4.0	4.5	4.2
8	GA	1.00	B	B	3.5	4.0	3.5	3.0	4.5	3.7
9	SF	1.00	B	B	3.0	3.5	3.0	3.5	3.5	3.3
10	AB	1.00	B	B	3.5	3.5	3.5	3.5	3.5	3.5
11	RH	1.00	B	B	3.5	3.0	3.5	2.5	3.0	3.1
12	GA	1.00	C	B	3.5	4.5	4.0	4.0	4.0	4.0
13	SF	0.00	A	A	3.5	4.0	5.0	3.0	5.0	4.1
14	AB	0.00	D	D	3.5	3.5	4.5	3.5	4.5	3.9
15	RH	0.33	B	B	3.0	4.0	2.0	3.5	2.5	3.0
16	GA	0.50	E	C	4.0	3.0	4.0	3.0	4.0	3.6
17	SF	1.00	A	A	4.0	4.0	4.5	4.0	4.0	4.1
18	AB	0.00	C	C	3.5	3.0	4.5	3.0	4.5	3.7
19	RH	0.50	C	B	3.0	4.0	3.0	3.0	3.0	3.2
20	GA	0.08	D	B	4.0	4.0	4.0	3.5	4.0	3.9
21	SF	1.00	A	A	3.5	2.5	3.5	3.5	3.5	3.3
22	AB	0.00	A	A	2.5	3.0	2.5	3.0	2.5	2.7
23	RH	0.50	B	A	3.5	2.5	4.5	3.0	4.0	3.5
24	GA	1.00	B	A	3.5	4.0	3.0	4.0	3.0	3.5
25	SF	0.00	A	A	3.0	2.5	3.5	3.0	3.0	3.0
26	AB	0.50	B	A	2.5	2.0	3.0	2.5	2.0	2.4
27	RH	0.33	B	A	3.0	3.5	4.0	3.5	3.5	3.5
28	GA	0.67	C	A	3.0	2.5	2.5	3.0	2.5	2.7
29	SF	0.00	A	A	3.5	4.0	4.0	4.0	4.0	3.9
30	AB	0.00	B	A	4.5	3.5	3.5	5.0	3.5	4.0
31	RH	0.25	B	A	2.5	2.5	2.0	2.5	2.5	2.4
32	GA	0.00	B	A	2.5	2.5	4.0	2.5	3.5	3.0
33	SF	0.00	D	D	4.0	4.0	5.0	4.0	4.5	4.3
34	AB	1.00	B	B	3.0	3.0	3.0	3.5	3.0	3.1
35	RH	0.00	D	D	2.5	4.0	4.5	2.5	3.5	3.4
36	GA	1.00	C	A	3.0	4.0	4.0	4.5	3.5	3.8

positively valued the methodology (75.64 and 73.81 percent). They also liked working in groups (79.49 and 90.48 percent).

In general, the students of the second cohort showed a greater degree of agreement with each sentence in the questionnaire (except for item 11). In particular, they seemed to have a much greater interest in being allowed to define their own shape grammars (71.43 percent versus 48.72 percent in the first cohort). There is not enough evidence to state whether this increased satisfaction level or interest in shape grammars has been due to the possibility of a more interactive experience with the tool or whether it could be due to other factors (like for example, the students' motivation).

The analysis of the two free-text items in the questionnaire supports these conclusions. The results of both cohorts are quite consistent in which aspects they liked most: Table 2 shows that the most frequently mentioned *theme* was *Support for creativity* (55.56 and 44.19 percent). Specifically, the category *The tool provides good starting points* (30.37 percent in Group A and 22.10 percent in Group B). An example of a supporting quote for this category is: "The range of starting points and solutions provided by the tool". Also in this *theme*, a positive perceived aspect was *overcoming*

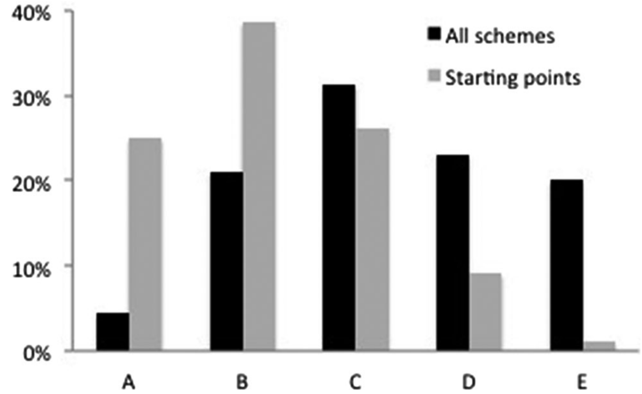

Fig. 6. Percentage of A-E in all schemas versus starting points (Group B).

preconceived solutions (8.89 and 8.14 percent in groups A and B, respectively). The instructors's feedback also pointed to a similar conclusion, when they said "However, the most interesting projects have emerged from accidental elements like the small annex spaces or errors". In fact, and according to their experience, "the designs of the clusters of schemas obtained using traditional methods are usually more rigid and less creative than the ones generated with the tool". Also, it seems that the students found that the tool helped designers to be more efficient (7.41 and 5.81 percent).

Following closely, the next most frequently mentioned *theme* is *Aspects relative to the quality of the solutions* (25.18 and 33.72 percent), and, in particular, *diversity* (19.26 percent in Group A and 18.61 percent in Group B). An example of a supporting quote for this category is "The tool provides numerous alternatives". In this sense the tutors declared that "the program generates such a wide variety of schemas that accidents occurred randomly, creating proposals that at first sight could be considered as undesirable forms, but in the end were used to generate the most interesting projects".

All in all, it seems that both cohorts of students found that the starting points provided by the tool were useful in the design process, interesting, diverse, and helped them to trigger their creativity when designing residential projects. As for the tutors, they valued positively the randomness and diversity of the solution provided by the tool, which encouraged analysis and reflection and expedited the design.

There are also some aspects to be improved, according to the students' opinions (as expressed in Table 3), and according to the instructors participating in the study.

The most frequently mentioned *theme* in both cohorts was *Aspects relative to the distribution of spaces* (48.04 and 42.37 percent), and specifically the categories *overlapping of specialized spaces* (10.78 and 13.56 percent); *better location of the front door* (11.76 and 13.56 percent) and *residual spaces* (9.81 and 10.17 percent). Examples of supporting quotes are "Spaces should not overlap, especially when there is free space"; "The front door does not need to be near the kitchen" and "Avoid residual spaces".

However, the next most frequently mentioned *theme* is different for the two cohorts: while for Group A it was *Aspects related to shape grammars* (21.57 percent), Group B students seemed to be more concerned about *Aspects relative the software tool* (28.82 percent). Group A students demanded additional criteria in the shape grammars (19.61 percent) while Group B students considered that the software tool should provide more editing capabilities (18.65 percent).

Finally the *theme Aspects related to solutions* was mentioned by both groups of students (17.65 percent in Group A and 22.03 percent in Group B). Both groups demanded more variety of solutions. Some students in Group A complained about the excess of randomness (8.82 percent), while some students in Group B said the tool should not present absurd solutions (6.78 percent).

The fact that students in the first cohort were critical with some aspects concerning the basic house schemas provided by the tool suggested that an increased user control could help to produce starting points which are more suitable for user's needs. To accommodate this, we implemented the second prototype, which allowed for more user interaction. However, students of the second cohort did not express clear preference for either of the two working modes.

The instructors also thought that the tool was at an early stage and should include additional architectural criteria to be used in professional practice. However, they found that its use in the academic context was more than adequate.

Next we will present some conclusions regarding the projects presented by students of the second cohort. These results provide answers for research questions 2 and 3. Though the sample size is small (36 projects, developed by 50 students), it is in the range of sample sizes for similar studies that we have discussed in the literature review section (10 to 62 students), so we think it is enough to at least obtain some evidence for the two additional research questions specifically developed for this second experiment.

Research question 2. Do students have any preferences for interactive or automatic solutions?

Here we discuss the preferences of the students for automatic or interactive solutions. The multiple choice question included in the questionnaire for Group B was aimed to understand the students' thoughts on this question. As explained in Section 3, when the students were asked about the preferences concerning the interactive or automatic modes, no clear pattern could be detected: roughly one-third of the students preferred one of the modes over the other (29 and 38 percent for the automatic and interactive, respectively), while about another third (33 percent) considered that each mode had a different functionality.

Items 13 to 15 of the questionnaire seemed to confirm this belief. Most of the 42 students who completed the questionnaire liked to be able to combine interactiveness with randomness (75.19 percent), but also to keep a certain degree of control with the interactive mode (92.86 percent). More interestingly, the greater majority of the students agreed that, in the automatic mode, randomness can help trigger the creative leap (90.48 percent).

This is also confirmed by the students' behavior: from the 88 selected schemas, 52.27 percent were interactive and 47.73 percent were automatic; and from the 36 projects, 38.89 percent were entirely based on automatic schemas, 30.56 percent were entirely based on interactive schemas, and 30.56 percent were based on a mixture of interactive and automated schemas. Note that the students' behavior when choosing cells to develop their projects is quite consistent with their thoughts as expressed in the multiple choice item, as approximately one-third of the projects were based on each option.

TABLE 5
Quality of Schemas versus Quality of Designs (Group B)

q_w	#	$\overline{Q}$	$\overline{Q_c}$
A	7	3.36	3.21
B	15	3.41	3.30
C	9	3.56	3.44
D	4	3.88	3.50
E	1	3.60	4.00

Research question 3. Does the nature (automatic/interactive) or quality of the initial seed have any impact on the perceived quality of the final residential projects presented by the students?

With respect to the nature of the initial seed, we have computed the coefficient of correlation between I and Q and it is $c = 0.04$, that is, practically zero. Moreover, if we compute the average quality for projects entirely based on interactive schemas, for projects entirely based on automated schemas, and for projects based on a mix of interactive and automated schemas, we obtain values of 3.6, 3.2 and 3.6, respectively. These results suggest that allowing users certain degree of control over the generation of starting points has not had any positive impact on the quality of the final projects. Therefore perhaps interactivity is not really needed to generate useful starting points that produce high quality designs.

Regarding the relationship between the quality of the starting points and the quality of the final designs, it can be discussed in the light of the data presented in Table 5. Each row summarizes the data for all projects whose worst schema used was of quality q_w. Each row shows the number of projects (#), their average quality as assessed by the instructors ($\overline{Q}$) and their average grading concerning creativity ($\overline{Q_c}$).

It can be seen that both global quality and creativity are better for projects generated from D and E starting points. To this respect, instructors believed that somehow worse starting points triggered the creativity of the students, who were able to generate good quality designs. It might also be the case that the more motivated students were willing to explore the complete range of solutions provided by the tool, and their intrinsic motivation also made them produce the best projects. All in one, it seems that, at least in this experiment, the quality of the starting point was not significant for the generation of good designs

It is also interesting to explore the quality of the seeds that the students have selected for each project. Fig. 7 shows the number of starting points of each perceived quality that has been selected in each type of project.

It can be seen that for the single-family houses, the students show a preference for the best starting points (with few exceptions), while for groupings they are willing to use all types of initial solutions. Probably for groupings they are considering other selection criteria (for example looking for an easy-to-group exterior shape), as it is not difficult to make the necessary adjustments to improve the quality of an initial seed Another possible explanation was given by the students themselves in the classroom presentations. In one of the student's own words:

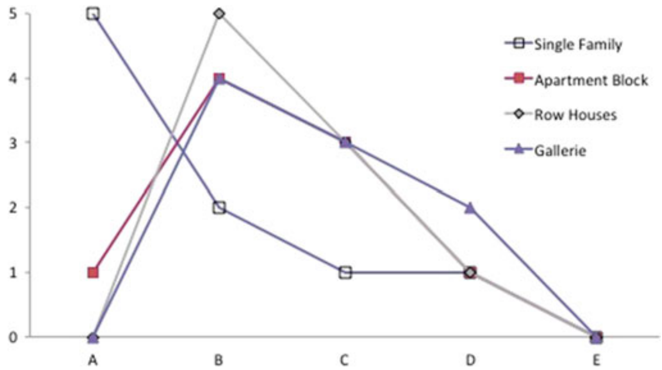

Fig. 7. Number of A-E in all schemas versus starting points (Group B).

"Some starting points that initially were labelled as problematic were selected for grouping projects, because usually in such projects there is a need to sacrifice the perfection of each single-family house. Also, the use of twists and irregularities in the initial seeds has given birth to creative projects. And it was useful to have such a large repertoire of initial seeds to choose from, because you could always find a cell that fitted in your design".

6 Conclusion

In this paper we have described the implementation and evaluation of two prototypes of an educational software tool (BH-ShaDe) specifically designed to assist architecture students in the early stages of a design, by providing them with starting points for the design of residential projects. The tool is based on intelligent techniques, namely shape grammars and reinforcement learning.

In order to determine the usefulness of the tool we have conducted an experiment in the Architecture School of the University of Málaga. The evaluation with 120 students of two different cohorts of successive academic years has shown that the tool was useful to students, by providing a diverse range of starting points that helped them as sources of inspiration and expedited the design.

The findings obtained in the second experiment of our study could also be of interest when developing computer tools to support design teaching and learning, though in this case the sample size is small (36 projects, developed by 50 students) so probably the results are not as conclusive. A first reflection can be made about where we must concentrate our efforts. As we have seen, it does not matter whether schemas are provided "as such" or they are the result of an interactive process; no clear preference is expressed by the students and no clear difference can perceived in the results.

Another important point for the design of educational software tools is that starting points do not need to be "perfect" to be useful; in fact, some students mentioned in their presentations that the "irregularities" of some basic house schemas triggered their creativity and they were able to use them productively. The instructors also shared a similar position with respect to this aspect. It was also confirmed in the evaluation of the projects, which showed that both global quality and creativity are better for projects generated from D and E starting points. Therefore, even if some students had preferred "more elaborated" starting points, it would be in opposition to the positive effects of ambiguities and imperfections, as perceived by both the students and the instructors as a trigger for creative work.

Additionally, the initial number of starting points provided by a tool should be, in our opinion, as great as possible, but then controlled by a conscious evaluation and assessment of the available alternatives. The initial solutions can then be modified (if needed) to suit other criteria. As several students said, one of the strong points of the software tool presented here is the number and diversity of initial seeds.

Acknowledgments

The authors would like to thank the instructors involved in the evaluation experiments for their enthusiasm and availability to cooperate with them in this study, and the 128 students for their hard work and for their amazing project proposals. The first author would also like to thank Dr. Rose Luckin and London Knowledge Lab for hosting her in Autumn 2014 and providing a rich research environment to work in this study. This work has been partially supported by grant TIN2009-14179, Plan Nacional de I+D+i, Gobierno de España and by PIE 13-052, Universidad de Málaga. Manuela Ruiz-Montiel is funded by the Spanish Ministry of Education through the National F.P.U. Program.

References

[1] H. Simon, "The structure of ill-structured problems," *Artif. Intell.*, vol. 4, pp. 181–201, 1973.
[2] A. Mitrovic and A. Weerasinghe, "Revisiting ill-definedness and the consequences for itss. Artificial intelligence in education: Building learning systems that care: From knowledge representation to affective modeling," in *Proc. 14th Int. Conf. Artif. Intell. Edu.*, 2009, pp. 375–382.
[3] C. Lynch, K. Ashley, V. Aleven and N. Pinkwart, "Defining ill-defined domains; A literature survey," in *Proc. Intell. Tutoring Syst. Ill-Defined Domains Workshop*, 2006, pp. 1–10.
[4] H. A. Simon, Information-processing Theory Human Problem Solving. In: Estes, W.K. (ed.) *Handbook of Learning and Cognitive Processes*. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, vol. 5, 1978, pp. 271–295.
[5] N.T. Le, F. Loll, and N. Pinkwart, "Operationalizing the continuum between well-defined and ill-defined problems for educational technology," *IEEE Trans. Learning Tech.*, vol. 6, no. 3, pp. 258–270, Jul.-Sep. 2013.
[6] P. Fournier-Viger, R. Nkambou, and E. Mephu-Nguifo, "Building intelligent tutoring systems for ill-defined domains," *Advances Intell. Tutoring Syst. Stud. Comput. Intell.*, vol. 308, pp. 81–101, 2010.
[7] P. Fournier-Viger, R. Nkanmbou, E. Mephu-Nguifo, A. Mayers and U. Faghihi, "A multiparadigm intelligent tutoring system for robotic arm training," *IEEE Trans. Learning Tech.*, vol. 6, no. 4, pp. 258–270, Oct.-Dec. 2013.
[8] C. M. Eckert, M. K. Stacey, and P. J. Clarkson, "Algorithms and inspirations: creative reuse of design experience," in *Proc. Greenwich Symp.: Digit. Creativity*, 2000, pp. 1–10.
[9] A. Andia, "Reconstructing the effects of computers on practice and education during the past three decades," *J. Archit. Edu.*, vol. 56, pp. 7–13, 2002.
[10] R. Hanna and T. Barber, "An inquiry into computers in design: Attitudes before attitudes after," *Design Stud.*, vol. 22, no. 3, pp. 255–281, 2001.
[11] K. Terzidis, *Algorithmic Architecture*. Oxford, U.K.: Architectural Press, 2006.
[12] A. Heylighen and H. Neuckermans, "Baptism of fire of a Web-based design assistant," in *Computer Aided Architectural Design Futures 2001*. Berlin, Germany: Springer-Verlag, 2001, pp. 111–124.
[13] C. Elsen and P. Leclercq, "SketSha" - The sketch power to support collaborative design," in *Proc. 5th Int. Conf. Cooperative Design Vis. Eng.*, 2008, pp. 20–27.

[14] X. Wei, D. Weng, Y. Liu and Y. Wang, "Teaching based on augmented reality for a technical creative design course," *Comput. & Edu.* 2014. DOI: 10.1016/j.compedu.2014.10.017.

[15] H. Y. Sunga, G. J. Hwangb, S. Y. Liub and I. J. Chiuc, "A prompt-based annotation approach to conducting mobile learning activities for architecture design courses," *Comput. & Edu.*, vol. 76, pp. 80–90, 2014.

[16] H. Liu and M. X. Tang, "Evolutionary design in a multi-agent design environment," *Appl. Soft Comput.*, vol. 6, pp. 207–220, 2006.

[17] L. Qjan, and J. Gero, "Function-behavior-structure paths and their role in analogy-based design," *Artif. Intell. Eng. Design, Anal. Manuf.*, vol. 10, pp. 289–312, 1996.

[18] A. Chackrabarti, P. Sarkar, B. Leelavathamma, and B. S. Nataraju, "A functional representation for aiding biomimetic and artificial inspiration of new ideas," *Artif. Intell. Eng. Design*, vol. 19, no. 2, pp. 113–132, 2005.

[19] G. Stiny, "Introduction to shape and shape grammars," *Environment Planning B: Planning Design*, vol. 7, no. 3, pp. 343–351, 1980.

[20] D. A. Al-kazzak and A. H. Bridges, "A framework for adaptation in shape grammars," *Design Stud.*, vol. 33, no. 4, pp. 342–356, 2012.

[21] G. Cagdas, "A shape grammar model for designing row-houses," *Design Stud.*, vol. 17, pp. 35–51, 1996.

[22] M.C. McGill, "A visual approach for exploring computational design" MSc Thesis, Department of Archit., Massachusetts Inst. Technol., Cambridge, MA, USA, 2001.

[23] M. V. Belmonte, E. Millán, M. Ruiz-Montiel, R. Badillo, J. Boned, L. Mandow, and J. L. Pérez-de-la-Cruz, "Randomness and control in design processes: An empirical study with architecture students," *Design Stud.*, vol. 35, no. 4, pp. 392–411, 2014.

[24] I. Iordanova, T. Tidafi and G. Paoli,, "Is a digital model worth a thousand pictures?," in *Computer Aided Architectural Design Futures 2001*. Berlin, Germany: Springer-Verlag, 2007, pp. 85–98.

[25] G. Goldschmidt and M. Smolkov, "Variances in the impact of visual stimuli on design problem performance," *Design Stud.*, vol. 27, pp. 549–569, 2006.

[26] H. Casakin, and G. Goldschmidt, "Reasoning by visual analogy in design problem solving: the role of guidance," *J. Planning Design: Environment Planning B*, vol. 27, pp. 105–119, 2000.

[27] H. Demirkan, "From theory to practice: 39 opinions," In A. Williams. M. J. Ostwald and H. H. Askland, Eds. *Creativity, Design and Education: Theories, Positions Challenges*. Strawberry Hills, New South Wales, Australia: Australian Learning & Teaching Council, 2010, pp. 56–59.

[28] A. M. Hill, "Problem solving in real-life contexts: Alternatives for design in technology education," *Int. J. Technol. Design Edu.*, vol. 8, no. 3, pp. 203–220, 1998.

[29] C. J. Atman, M. E. Cardella, J. Turns and R. Adams, "Comparing freshman and senior engineering design processes: An in-depth follow-up study," *Design Stud.*, vol. 26, no. 4, pp. 325–357, 2005.

[30] A. Heylighen, P. Deisz and I. M. Verstijnen, "Less is more original?," *Design Stud.*, vol. 28, no. 5, pp. 499–512, 2007.

[31] G. Fricke, "Successful approaches in dealing with differently precise design problems," *Design Stud.*, vol. 20, no. 5, pp. 417–429, 1999.

[32] M. Ruiz-Montiel, M.V. Belmonte, J. Boned, L. Mandow, E. Millán and J. L. Pérez-de-la-Cruz, "Layered shape grammars," *Comput.-Aided Design*, vol. 56, pp. 104–119, 2014.

[33] M. Ruiz-Montiel, F. López, E. Millán, M. V. Belmonte and J. L. Pérez-de-la-Cruz. (2012). "BH-ShaDe user's manual," [Online]. Available: http://tinyurl.com/nwfs7yx

[34] Montaner Muxí arquitectes, "Propuesta de nueva normativa de viviendas," Dirección Gen. de Ordenación del Territorio, Junta de Andalucía, Tech. Rep., 2008.

[35] M. Ruiz-Montiel, J. Boned, J. Gavilanes, E. Jiménez, L. Mandow and J. L. Pérez-de-la-Cruz, "Design with shape grammars and reinforcement learning," *Advanced Eng. Inform.*, vol. 27, no. 2, pp. 230–245, 2013.

[36] L. Cohen, L. Manion and K. Morrison, *Research Methods in Education*. Evanston, IL, USA: Routledge, 2011.

[37] S. Jamieson, "Likert scales: How to (ab)use them," *Med. Edu.*, vol. 38, no. 12, pp. 1212–1218, 2004.

[38] A. Strauss and I. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* 3rd ed. Newbury Park, CA, USA: Sage, 2008.

[39] B. G. Glaser and A. L. Strauss, *Discovery Grounded Theory*. Chicago, IL, USA: Aldine Transaction, 1967.

[40] K. Dorst and N. Cross, "Creativity in the design process: Co-evolution of problem-solution," *Design Stud.*, vol. 22, pp. 425–437, 2001.

[41] H. Janson and U. Olsson," A Measure of Agreement for Interval or Nominal Multivariate Observations," *Edu. Psychological Meas.*, vol. 6, no. 2, pp. 277–289, 2001.

Eva Millán received the PhD degree in computer science from the University of Málaga in 2000. She is currently an associate professor at the same university. Her research interest include intelligent tutoring systems and educational software tools, user and student modeling, Bayesian networks and, more recently, applications of AI techniques to the development of software tools for architectural design teaching and learning.

María Victoria Belmonte received the PhD degree in computer science from the University of Málaga in 1992. She is currently an associate professor with the Departamento de Lenguajes y Ciencias de la Computación at the same university. Her current research interests include multi-agent systems and, more recently, tutoring systems for architecture and artificial architecture (computer aided design based on shape grammars).

Manuela Ruiz-Montiel received the MSc degree in computer science from the University of Málaga in 2010, where she is currently working toward the PhD degree. Her research is focused on developing new AI techniques, based on search algorithms and reinforcement learning, for the implementation of computational design methods based on shape grammars. She has applied these methodologies to several architecture related problems, like housing and mega-structure design.

Juan Gavilanes received the PhD degree in architecture from the Polytechnic University of Madrid in 2012. He is a currently a professional architect and also an assistant professor at the University of Málaga. His research interest include applications of AI to architectural design problems. Some of his constructed works have received important awards, from year 1996 to the present days.

José Luis Pérez de la Cruz received the PhD degree in engineering from the Polytechnic University of Madrid in 1991. He is currently a professor of computer science and artificial intelligence at the University of Málaga. His research interests focuses on the design of new computational techniques for AI problems and their applications to different contexts, including engineering, education, and architecture.