Chapter 6

Experiment and Assessment

The preceding chapters have served to situate and describe a program that endeavors to answer the research question posed in Chapter 1: *Can a computer application for the design and construction of mechanical automata accelerate the acquisition of mechanical reasoning and spatial cognition in children or increase children's proficiency in these areas?*

Chapter 2 reviewed the literature that focuses on mechanical reasoning and spatial cognition and found many examples where machines were used to study these abilities, but few instances where the study of machines was being used explicitly to affect changes in mechanical reasoning. This information was used to devise an approach for constructing reasoning through the design and creation of small machines. Chapter 3 examined a variety of tools that could be used to create machines in both virtual and physical environments and looked for attributes that could facilitate the design and construction of small machines for novices. All of the tools reviewed suffered from one or more flaws that limited their usefulness for children learning to create small machines. These flaws were addressed in the design requirements for the MachineShop system to help ensure that it would provide the support that users would need. Chapter 4 presented mechanical automata as an appropriate and interesting domain in which to create machines. In particular, the characteristics that define contemporary automata were examined and shown to be salient for novice machine designers and builders. Chapter 5 described the software and hardware components of the MachineShop system that resulted from the analysis of the material from the chapters that preceded it.

This chapter describes the experimental protocol devised to evaluate the use of the MachineShop system as it applies to the thesis question. It begins by examining the experimental method with a focus on the intended user population, the form of the user testing, and the plan for assessment. It continues with an exploration of the results obtained from the testing and concludes with a discussion of the methods and results and considers to what extent it is possible to answer the thesis question.

6.1 The Experimental Method

From its earliest incarnation, this research was intended to introduce children to the work practices and techniques of adult designers and engineers. It remains a firmly held belief that by presenting children with an authentic look at professional practice they will gain a clearer understanding of what it means to pursue that profession as a vocational objective. This objective constrained many choices for the design of the experimental method, from the domain to the users and from the tools to the process.

6.1.1 The Users

MachineShop was originally conceived as a tool for children as young as eight years old, but as the realities of creating a system like this became apparent, this was shifted upward somewhat, resulting in a target group aged eleven to thirteen years. It was further assumed that children using this tool (both during the testing and at other times) would not be familiar with contemporary automata or any other mechanical contrivances of this general type. Beginning MachineShop users were expected to have some experience building objects (with construction kits being the most likely method) but were not expected to be familiar with the tools or materials they would use in the construction of automata.

6.1.1.1 Recruiting Users

The users who participated in this research were recruited from two pools. The first consisted of the children of faculty and graduate students in the Department of Computer Science at the University of Colorado at Boulder. From an initial group of five interested children, three participated in the testing. A middle school science teacher in Boulder provided suggestions for students that formed the second pool. Five children expressed an interest and three of these participated in the testing. One user was a member of both groups. All users were self-selected and participated with the encouragement and cooperation of their parents.

6.1.1.2 User Demographics

The users for this study were two girls and four boys aged 10 years and 6 months to 12 years and 1 month at the beginning of their involvement with the testing. This placed them in the lower half of the anticipated age range and in some cases below the lower bound initially intended. All of the children attended public schools in Boulder; two were fifth-grade students at an elementary school and four were sixth-grade students at a middle school. In addition to their schoolwork and participation in this study, all were engaged in extra-curricular activities including sports, music, theatre, academic clubs and groups, and organized non-competitive games. Table 6.1 provides some information about the users and their participation for comparison.

User	Age	Grade	Participation	Automata Built
Dylan	10 years, 6 months	6	37 weeks	2
Calum	11 years, 8 months	6	21 weeks	1
Sam	11 years, 5 months	6	25 weeks	2
Frank	12 years, 1 month	6	20 weeks	1
Abbie	10 years, 10 months	5	16 weeks	1
Iris	10 years, 11 months	5	13 weeks	1

Table 6.1: A summary of information about the test users in the study. Users are listed in the order in which they started participating and their ages at that time are also shown. Pseudonyms are used here and in Chapter 7.

6.1.1.3 Prior Experience

All of the users did have prior experience with designing and building physical objects. This most often involved the use of construction kit toys of the type discussed in Section 3.3. While Kinex and Zometool [134] were prominent in this collection, all users had the largest part of their experience using products from the LEGO Group, including building moving constructions from Technic components and making robots using Mindstorms kits. In many cases the activities engaged in by the users with construction kits were social activities in which friends, siblings, and often parents participated. Two of the users had also built plastic model kits (airplanes and cars) while one had worked with her father in constructing a wooden whirligig for their garden. Another user had built a railroad locomotive model from found materials of wood, plastic, and metal. None of the users were initially familiar with mechanical automata of any form or with any related devices other than music boxes.

The tools most familiar to the users included screwdrivers, pliers, hammers and paint-brushes. Their use of drills, knives, and saws was more limited and almost universally undertaken while being supervised by an adult. The use of power tools other than electric drills was almost nonexistent. Users also had varying levels of experience using adhesives and paint. None of the users possessed the range of skills and knowledge necessary to undertake the construction of an automaton and that provided a useful baseline for evaluating the results of the tests.

6.1.2 The Experimental Procedure

Two potential approaches to the user testing were considered. Devising an educational curriculum around the design and construction of automata has a certain pedagogical charm and is the method by which thousands of children (see Section 2.3.3) are introduced to these types of machines. But it also poses significant difficulties for research designed to fit within the scope of a dissertation. First is the problem of creating an epistemologically sound program that can be used effectively with entire classrooms of children over extended sequences of short class periods. Second is finding one or more classroom teachers who are willing to expend the additional resources necessary over a full school year (the length of the longest cases reported here) to assist in making such a program work, particularly in light of recent trends toward assessment based public education [21, 124][1] . Third is the need to recruit additional researchers to interact with and assist students over the course of the testing. Fourth is the massive task of relaying the observations of a group of researchers to the primary investigator and making sense of it all. For these reasons this approach was set aside.

The use of case studies was the other approach considered. It promised a more tractable method with a higher probability that the primary (and in this case sole) investigator would not overlook important occurrences during, or results from, the user sessions. It would support an opportunity for the interaction between the test users and the investigator to be more like an apprenticeship than might be the case in other approaches. This was desirable because it allowed the test users a great deal of freedom in the manner in which they approached the work and encouraged them to learn from their choices and actions while still maintaining the presence of a collaborator who could provide appropriate levels of assistance when necessary. It additionally offered a more relaxed environment in which to work without the distractions created by two dozen children engaged simultaneously in a potentially chaotic activity. Case studies also have their shortcomings, most notably the small quantitative data sets they can provide which makes it difficult to extrapolate results to larger populations. Still, it offers an approach both practical and time honored [126] and, because it has proven successful in similar research efforts (see [32] for one example) it became the chosen method.

6.1.2.1 The Design Process

The work that I envisioned the children doing was derived from my experiences working in engineering organizations and was intended to distill the essence of engineering process into

[1] Many innovative teachers are finding it increasingly difficult to convince administrators to accept novel curricula with no proven track record. This is particularly true where schools and districts receive financial inducements to perform well on the standardized assessments.

a form that children could adopt. In an ideal (for both the business and their customers) world, the procedures used to create a new product would approximately be:

(1) Determine a need. Profit is the motivating factor in product design for a company and often a tremendous amount of effort is expended in doing the research necessary to evaluate the money making potential of a new product before any further steps are taken.

(2) Devise and evaluate solutions. Once a need has been identified as a potential source of revenue, ideas are floated for possible solutions. These are evaluated and modified until one is determined to be best from a number of standpoints; satisfying the need, profit potential, support of core business goals, and many others.

(3) Propose and evaluate designs based on chosen solution. Designs for products that best embody the solution are proposed. Competing designs are compared and evaluated and the best features from each are combined, if possible, into a final design.

(4) Iteratively prototype and refine the chosen design. A prototype based on the final design is created and evaluated. Flaws in performance and difficulties in manufacturing are identified and corrected over a series of successive prototypes.

(5) Document the final design. Documentation is created that describes how the product is to be manufactured, how it is to be used, what its operational aesthetic characteristics are, how much it will cost to produce and procure, and so forth.

(6) Produce the product. At this point raw materials are collected, tooling is produced, staff or subcontractors are employed, and warehousing and distribution arrangements are made.

(7) Market the new product. Advertisements are created and placed with media organizations that have the broadest contact with the intended user population. Orders arrive,

providing the business with capital to begin the process over again for a new product.

These steps are very coarse-grained and do not reflect most of the tasks necessary to make them happen. It is also clear that several of these steps have no direct connection to the engineering portions of the process. By culling those that do and adapting them to the creation of contemporary automata, the procedures derived for the children in these trials took on the following form:

(1) Envision an automaton that you would like to build. It should be inspired by something personal: an event, interest, question, statement, or idea that has meaning for you. Think about the story you would like to tell and how your automaton would tell that story. Determine what figures, props, and settings would be needed to tell your story.

(2) Decide what actions and behaviors your automaton may need to tell the story you have chosen. Don't worry about how these actions and behaviors will be implemented, but instead work at refining what you want your automaton to do.

(3) Determine ways in which the mechanism for your automaton can provide the actions and behaviors you've decided to use. Don't settle for the first idea you conceive but explore alternatives. Some of them may be even better at doing what you want.

(4) Make a prototype of your design from materials that are easy to work with. Don't worry if the prototype doesn't work like you want the finished automaton to. Its purpose is to help you learn about your design and how it can be made better. Once you understand where your design still needs work, come up with fixes and make them part of a new prototype.

(5) When you are sure that your design tells the story you want and works the way it should, make the pieces that you will need for the finished automaton and put it together. Decorate and embellish it as you see fit.

(6) Enjoy the results of your hard work. Think about what you would like to build next.

This process clearly owes its form to the commercial process, but is motivated by the child's desire to create an object that has personal significance rather than a profit. While there was some variation from user to user, these steps were followed consistently with support being provided to the users in the smallest amounts necessary and only when essential to allow users to successfully complete each step. Since users worked one-on-one with the researcher, it was possible to tailor this level of scaffolding to the requirements of each user. Some tasks, such as sketching and iterative prototyping, had been identified early on as important components of the scaffolding that would move from being initiated by the researcher during early phases in the testing to becoming integral and natural components of the task as work progressed.

Testing began by introducing each user to the domain of contemporary automata. This was accomplished by showing them photographs and videos of professionally created examples, giving them access to educational materials such as posters and books, and by providing them with the chance to play with a small number of example automata that I had previously created using early versions of the MachineShop system. The user and the researcher spent some time talking about the various features of all the examples, emphasizing the stories that they told and the behaviors they exhibited while limiting the discussion of how the actions and behaviors were obtained. At the end of this first session, the users were asked to return for the next session with, at a minimum, an idea for and a sketch of the automaton that each wanted to create.

The requested sketch served three important functions. First, it provided the researcher with some understanding of what the user hoped to accomplish with respect to style and scope and provided a first instance of what was hoped to be a series of drawings that would document the design process. Second, it provided insight into the user's comfort level with the task. This was gauged by the level of detail included and whether the user had provided any hints at a possible mechanism. This latter seemed plausible in light of each user's exposure to the example automata in the introductory session given the ways in which the combination of figures and mechanism serve to define the behavior of an automaton. Third, asking the user to create a drawing at the very beginning of the testing helped integrate sketching into the entire process in a natural way. The role of drawings in these tests will be discussed further in Section 6.1.3.

The users were also sent home with a set of four pre-test problems[2] in order to get a sense of their initial mechanical reasoning ability. Each problem consisted of a drawing depicting the input and output elements of a mechanism with its workings not shown and were based on simple examples presented by Cox [23]. The input and output motions were described and the users were asked to add mechanical components to the drawing that they believed could account for the stated operation of the mechanism. Since the purpose of these exercises was to get a feel for the users' abilities to reason mechanically, no strict time limit was imposed, although parents were asked to see that no more than five minutes was spent on each one.

The next few sessions were spent moving from the initial automaton concept to a first candidate design. During these sessions the users were introduced to the MachineShop software and were given opportunities to revisit any of the introductory examples and materials that they wished. At this time they were also provided access to a number of simple mechanisms created with MachineShop but decoupled from automata (see Section 5.3.1) to assist with their assimilation of the behaviors provided by basic mechanical components. Decisions were made at this stage regarding the size of the automaton, the nature of the figures and setting, and defining a more precise specification for the behavior. It was also at this time that the design of the mechanism was explored, conceived, and refined. Sketches and crude approximations of the necessary pieces were made to give the users a better understanding of their designs and the ways in which the designs were changing.

When a design was considered to have met the design specifications, the MachineShop software was used to create the necessary mechanical components and produce fabrication files for the laser cutter (see Section 5.2.3). Designs for the figures were scanned and converted by the

[2] One user did not receive the pre-test as he had begun testing in the pilot phase, before the pre- and post-test materials had been completed. He did, however, complete the post-test problems.

researcher into fabrication files at this time. The laser cutter was used to produce initial pieces in cheap, readily available materials like corrugated cardboard that were then used to test for fit and function. These pieces were also used to determine the structure needed for the supporting framework for the mechanism and preliminary fabrication files for the structural supports were also created and mockups of these pieces cut. The assembling of these pieces provided insight into flaws in the design or behavioral concept and these were addressed by modifications to the design and the construction of another prototype. Users were constantly revisiting the example materials and using the software to make changes to their mechanism components.

When the design had progressed to the limit of the capabilities of the cardboard prototypes, components for a final assembly were fabricated. In some cases, this assembly helped identify problems which could not have been discovered with the previous prototypes (see Section 6.2) and further iterations were made. The creation of the final parts signaled entry into the final phase of the process. Assembling the components into an automaton was a task that was important not to rush. Users who had shown great patience during the prototyping were now anxious to see how closely their design lived up to their expectations. The researcher paid particular effort to ensuring that users developed good techniques for preparing components for assembly as well as for fitting and gluing, especially with regard to allowing glued components sufficient time for the bond to cure. In all cases, as components were cleaned[3] they were test fit. This provided assurance that the automaton would fit and function as desired as well as giving the users multiple chances to learn the proper orientation of components and to determine a satisfactory sequence for assembly before anything was permanently attached. After the automaton was assembled, it was tested for function. In some cases minor adjustments were required to provide smooth operation. At this point, some users chose to paint and decorate their automata, while others chose to leave the materials in their natural states.

[3] The laser cutter creates components by vaporizing material to create the kerf. When this is done to materials like cardboard or wood it leaves a charred surface on the component which, if not cleaned, can leave a residue in the mechanism as it is removed through the normal operation of the mechanism. Steel wool has been found to quickly remove this loose material without changing the profile of pieces which can easily happen when an abrasive such as sandpaper is used.

6.1.2.2 The User Testing

Users were asked to come to the laboratory once a week for sessions which lasted from 60 to 90 minutes. The longest participation by a user was for a period of 37 weeks during which two automata were constructed. The shortest participation was for 13 weeks during which a single automaton was constructed. With respect to the construction of a single automaton, the maximum time spent was 21 weeks, the minimum was 12 weeks, and the median time to create an automaton was 16.5 weeks (see Table 6.1). There were two primary factors that account for this rather wide range. The first was the complexity of the automaton, with two of the three most complex taking the longest times and the two simplest taking the shortest times. Interestingly, one of the most complicated automata was also among those taking the shortest times (see Section 7.2 for more detail). The second was the speed with which a user gained knowledge of the domain and the design process. This was reflected both in attention to detail and in the creation of designs that challenged their abilities. Rather than speeding up as skill increased, users tended to slow down. This is discussed further in Section 6.2.

6.1.3 Assessment

The assessment goals were twofold. First and foremost was the goal of answering the thesis question; to see if using the MachineShop system to create contemporary automata had an effect on the spatial cognition and mechanical reasoning abilities of the test users. Second, and no less important, was the goal of understanding how children approached the design process and to what extent their skills in this area improved. While these goals both evaluate highly correlated features of the user tests (one might even argue that they are merely different perspectives on the same feature), breaking them apart in this manner allows us to talk about aspects of the activity that are separate from its cognitive effects.

As previously mentioned, these tests were not conducted with large groups of subjects so no effort was made to provide statistical analysis of the results. Further, there was no attempt to compare the efficacy of competing approaches for acquisition and enhancement of mechanical thinking. Assessment of the users' performance was based instead on a number of qualitative measures suggested by the mechanical reasoning literature and described in Section 2.5. Because each user was interacting one-on-one with the researcher, it was possible to augment these metrics with observations made during test sessions. For example, there were occasions when a sudden insight was reflected in a facial expression or hand gesture or in sudden changes of mood or attitude. While not scientific criteria, these were nevertheless important pieces of the assessment puzzle and were significant harbingers of change.

6.2 Results

The results seen during the user testing follow closely from both the complexity of the automata created and from the length of participation in the tests. This is not surprising, as one would expect users to become more skilled and proficient with prolonged involvement. But this is a great oversimplification as the results showed far more in the way of differences than can be attributed solely to these measures. Although the number of users was small, the changes that were seen varied from negligible to striking. This section examines the changes that were observed in the users over the course of the testing with particular emphasis on noteworthy results. The interested reader will find more details in Appendix A.

6.2.1 Talking About Automata

None of the users had any exposure to mechanisms before starting these tests and knew little of the terminology of the domain beyond lever (from their introduction to simple machines in science class) and gear (from the incorrect name given the sprockets of a bicycle). Throughout the tests the researcher was careful to use the terminology that users were hoped to acquire and encouraged users to do the same. Initially, all of the users tried to apply the few terms they knew in as many situations as possible. This resulted in followers being called levers and cams being called gears (among others) for several weeks. Two users showed little progress in this ability over the course of the experiments and at the conclusion of testing were still using more generic terms (“cam” rather than ”eccentric cam”) even though the automata they were building used only eccentric cams and this was the term used exclusively by the researcher. The majority of users were able to pick up the descriptive terminology after several weeks and were using secondary terms (lobe, follower, tooth) correctly by the end of the testing. One user assimilated quite a large domain vocabulary in a very short period time about 12 weeks into his first automaton (this is discussed more in Section 7.1).

The ability to describe the workings of components and mechanisms using domain appropriate terms was a more challenging task for all users. Users who had shown improvement in their abilities to use appropriate descriptive terms with components were unable to consistently apply that skill when discussing components in assemblies and they often reverted to terms like “bump” and “stick” instead of “lobe” and “follower” which were terms they readily used when talking about the components in isolation. Even when describing the transmission of motion from component to component within a mechanism, users chose more general terms rather than domain terms that they knew. This was a surprising result. I had originally suspected that the use of descriptive terminology for components and more technical language to explain the function of mechanisms would change more or less concurrently, but in these tests the use of functional terminology trailed the use of descriptive terminology by a significant amount.

6.2.2 Using External Representations

The use of external representations during the design process took three forms. First were the example automata and simple movements (Section 5.3.1). These featured most prominently during the early sessions when users were first exploring their designs. In the beginning all users just played with these, but the majority quickly moved beyond simple play to an exploration of the features of the mechanisms and the motions that they provide. Most users returned to the simple movements over the course of the testing even after work was well under way on what would become the final designs. These objects supported the generation of new ideas for the design, were useful in acquiring domain knowledge, and were used as shared objects for the exchange of ideas between users and the researcher.

Next were the prototype objects. Somewhat interestingly, these objects were not initially given as much interest as the sample materials, possibly because their appearance was somewhat rough and they were not as robust. This tended to change over time however, and users began spending more time with, and effort on, their prototypes. Expectations had been that beyond being useful for evaluating fit and function in the evolving design, users would find that prototypes spurred new ideas and the acquisition of domain knowledge. This result was not seen to any substantial extent.

Finally were the drawings and sketches. Throughout the testing the researcher created, and encouraged users to create, drawings. These were used to codify ideas and stimulate new ones, record pertinent design details, and provide a visual journal of the evolution of the design. It had been expected that the users would create many sketches during the testing but none were spontaneously created. Only at the urging of or questioning by the researcher did users draw sketches. This was somewhat puzzling, and putting the question to the users revealed no good explanations for their reluctance. Perhaps the best explanation is that design process was an unfamiliar one and that, with more experience, they would understand the importance of drawing to design. As was shown in Section 2.5.2, drawings are very powerful forms of external representations and are capable of conveying large amounts of relevant information in simple forms. Goel [44] has shown how ill-structured representations in preliminary designs support the transformations necessary for design evolution. Suwa [75] suggests ways that hidden features in sketches can be discovered and used to advance the design. Clearly, this is a disappointing outcome and one that deserves to be investigated further.

6.2.3 Components, Motion, and Complexity

While described individually in Section 2.5, the abilities of mapping structure to motion, discrimination and integration, and dealing with complexity are intertwined in practice. This was obvious during the user testing when observed changes in one ability were reflected in similar changes in the others. By the end of the tests, all users could look at a drawing or photograph of a simple mechanism and describe the motion that would be produced. This was true even for users who showed little change on other metrics. When asked, users most often remarked that "playing" with the simple movements had helped them to better understand how the shapes of components and the ways in which they were connected to other components determined the behavior that would be seen.

But as mechanisms became more complex, the ability to discriminate became a limiting factor on the ability to map structure to motion. This was particularly true when the unknown mechanisms contained components that differed in significant ways from those with which the users were familiar. As an example, a poster from Cabaret Mechanical Theatre hangs on the wall of our laboratory near the laser cutter. It shows an automaton of two swimming fish with a mechanism composed of familiar components (levers), unfamiliar components (cams with unusual profiles), and familiar components used in unfamiliar ways (pinwheel gears that mesh with a rack). All of the users could identify the levers and gears, and most were able to identify the cams as cams. Two were able to fully describe the motion of the automaton by the end of the testing period. This is a wholly remarkable result given that the behavior of this automaton has puzzled many of the adults who have spent time with it.

Even less change was observed with respect to the users' abilities to integrate components into new mechanisms. The results of the post-tests showed uniformly that when given a mechanism design task with constraints the users preferred the components with which they were most comfortable, even if they were not good choices. It was often the case that if the chosen component was a poor pick, the user would attempt to modify it in some way in an effort to make it work better. This was observed with all users to some extent although there were variation across test problems for each user. It is noteworthy that there was little difference in the quality of the solutions between the pre- to post-tests although there were pronounced changes in the styles of the solutions.

Results were more promising for the users who created a second automaton. Although only two users are in this category, both of the later designs were more ambitious and complex than their first ones. In both cases they chose to work with components they had not previously used and to combine them into assemblies that they had not seen in any of the examples. In one of the cases, the user came to the researcher with a fully formed design for the mechanism and was able to produce a drawing of it for the researcher when asked.

6.2.4 Confidence Levels

With the exceptions of the users who showed the least change in all measures, there was a significant increase in the confidence of the users with respect to their abilities for creating automata. This was apparent in a number of ways: increasing enthusiasm, transfer of tasks from researcher to user[4] , content and intensity of conversations, and less fear of failure. This latter was particularly evident in users who were not only willing, but anxious to experiment with new ideas and concepts. Each of the two users who built a second automaton exhibited an increase in confidence (as well as in domain knowledge) with the more ambitious nature of the second design. While each was a more complicated design than the first, a similar amount of time was spent in developing both designs and in each case prototyping for the second design started earlier in design process than it had for the first design. As all users became more confident they began to ask questions that indicated they had developed understandings of the domain and the process that were more than superficial. In general, it seemed that the more confident users got more from the testing and had a better time as well.

6.2.5 MachineShop System Usage

As with most other tasks during the testing, the use of the MachineShop software varied over a considerable range. Two users spent only enough time with it to create the fabrication files

[4] As users gained confidence they spontaneously took over tasks from the researcher. These included using calipers and tape measures to determine size and distance, performing calculations about the physical and functional properties of their automata, and setup and operation of the laser cutter under the supervision of the researcher

they needed for their simple mechanisms. One user spent a considerable amount of time with the software, not only designing the components he needed, but also exploring most of the software components and by the end of the testing had become almost as familiar with its use as I am. The remaining users' time with the software was between these extremes. While none of the users found the software difficult to use, comfort with both the interface and operation of the software increased as more time was spent interacting with it.

Watching users interact with the MachineShop software provided valuable information about the design of each tool and their integration into the larger whole. The balance of this section presents a number of key observations made relating to the ways in which the users employed the software in the making of their automata and what these mean for the current form of the software. Remarks about each tool are presented in the order that the tools are described in Section 5.1.

The movement explorer (Section 5.1.2) received very little use, seldom more than one or two quick visits. In the context of the user testing this is not unexpected. All of the mechanisms available to users in the movement explorer were available to the test users in physical form since the simple movements present during testing were also used to create entries for the movement explorer. The simple movements were, however, used extensively both to shape the designs the users created and as aids in exploring and understanding the domain. Too little attention was paid to this tool to truly gauge either its usefulness or accessibility to users.

The component editors (Section 5.1.3) were far and away the most used tools during the testing. This was due mostly to the fact that, with the design computer and laser cutter being located in the same room, users could receive rapid feedback on their designs in the form of physical objects. With respect to the relationships between the software tools and the automata created by the system, that of the editors is unique. These are the tools that ultimately define, shape, and produce the finished products. For the users in these tests, who are just on the cusp of the period of formal operations [45], the opportunity to work with physical artifacts is still a very powerful knowledge building activity and the most intellectually productive sessions observed were when users were creating, examining, and manipulating components they had just designed. Users seemed to particularly appreciate the ability to view changes to component profiles in the editors in realtime, and it was not unusual to find users spending up to ten minutes creating and modifying components in the editors.

The libraries (Section 5.1.4) were not used at all during the testing. Several factors seem to be at work here. First, most users knew that they would only be creating a single automaton and that provided no incentive to save components for reuse or later modification. Components were saved only as fabrication files, and if there was a need for modification, users found it easy enough to create new components from scratch. Second, users were not working on their own computers, but rather on a shared computer in the lab. To most users this meant that they would not have access to the files after the testing period so creating those files offered no benefits. Third, there was no sense of community. While all users were aware that there were other participants, and in some cases knew who those others were, the constraints of the testing were such that users were discouraged from talking to and working with others. Saving files was seen as a first step in sharing designs, an activity that was proscribed in the testing.

6.3 Discussion

Beyond illuminating the changes in the spatial cognition and mechanical reasoning abilities of the users, this set of tests was a rich source of information about children and construction.

The users came to the testing with little experience with the type of work they would be doing and with many naïve assumptions about the process. Commonly, users expected to design and construct an automaton in just a few weeks but most realized this error early on and came to understand that by requiring more effort, the process would be more rewarding. During this perceptual shift, users started to understand that there were a variety of materials available to support their efforts. While all users had been introduced to these materials at the beginning of the testing, users did not seek them out until they were needed. It had been assumed that introducing users to resources and examples in both the lab and on the World Wide Web would stimulate an interest in the domain but in practice it was the need to progress beyond plateaus in the design process that stimulated that interest. When presented with the opportunity to use either a virtual resource (on the web or in the MachineShop software) or a physical resource, users overwhelmingly chose to pick up and manipulate the objects. A majority of sessions contained one or more uses of these resources, even after designs were completed and construction was underway, indicating that there was more to this activity than the simple acquisition of information necessary for the task at hand.

Most users showed similar behavior towards the creation of prototypes and spent more time evaluating and modifying them than might have been required. More than any other, this step in the design process was essential to the successful completion of the automaton and the user's satisfaction with it. Prototyping allowed users to work through difficult design problems, notably in the lion (Section B.3), the commuters (Section B.2), and the evil cat (Section B.5). While formal prototyping was a new task to users[5] most discovered its utility.

In these tests, prototypes were created in corrugated cardboard whenever possible. Cardboard is readily available and often can be acquired for free in the form of discarded packaging from businesses. These boxes are often large enough to duplicate the maximum material size that the laser cutter can accommodate (12 by 24 inches for the machine in our lab) and most desired material thicknesses can be obtained by laminating. Pairing highly available materials with the rapid fabrication capabilities of the laser cutter makes the prototyping step even more powerful. Without this ability it would be necessary to assess design changes through calculation or speculation before creating a modified prototype. With this ability, the quality of a modification can be assessed directly, further speeding the process. During the testing this allowed users to make incremental changes without regard to the number of component or support structure modifications that were required. The laser cutter[6] offered the further advantage of using the same

---

[5] All of the children had created prototypes in their previous activities but were unaware of it. In particular, all had used LEGOs to design and build and, when asked, admitted that this activity involved making an object and modifying it several (sometimes many) times before considering it finished. The design of LEGO components offers the affordance to build in this manner and none of the users saw the parallels across domains.

[6] The laser cutter is a fascinating device to almost everyone who sees it. The users were no different, and their

fabrication files for both prototype and final components, needing only changes in speed and power settings to work with different materials.

The creation of cardboard prototypes was not without its limitations. Problems most often arose because of the differences in structural properties between the materials used for the automata (most often wood and plastic in these examples) and cardboard. The wood and plastic employed were far more rigid for a given thickness than cardboard. This was especially evident in larger pieces, particularly structural supports. In the prototypes it was necessary to use gussets and stiffening ribs to eliminate unwanted movement of these parts and mechanical components such as pinwheel gears could not be made functional. Wood and plastic components are heavier than their cardboard counterparts. This makes mechanical components like cam followers and pawls that depend on the force of gravity for operation less likely to function consistently when made from cardboard. Wood and plastic components have different textures than cardboard which also affect the operating characteristics of the mechanism. This is particularly conspicuous when the cut edge of a cardboard component exposes its corrugations. These corrugations leave the edge surface uneven with respect to strength, stiffness, regularity of the surface. This in turn makes it difficult to ascertain how motion will be transferred between components that contact one another along their edges. When these conditions made it impossible to proceed further with refinement of the design it became necessary to start creating prototype components in the final materials, although in all automata if this was needed it was very close to the completion of prototyping.

Another interesting observation from the testing was the difficulty that the users had in thinking about size in the context of their automata. In nearly half of the cases, there were significant differences between the initial stated size and the size of the completed automata. In some cases, the initial size chosen by the user was half that of the finished automaton. This problem manifested itself in two ways. The first was a problem of scale. Ten and eleven year old children

---

captivation never ended. Some learned to configure the machine for cutting (always done with supervision from the researcher) although the constraints of the testing prevented them from learning everything necessary to fully utilize it.

(and many adults it should be added) have great difficulty in estimating size either accurately or consistently. This led to problems when the initial placement of components in a mechanism prevented them from aligning properly with other components or with their connection points on the moving parts of the figures. The second difficulty came when users were forced to identify the source of mechanism malfunctions. These were most often due to misalignments or improperly supported components. For example, shafts in gear trains must be properly spaced (or the gears will not mesh correctly) and they must be supported so that they run true without wobble or excessive endplay.

Finally, it should be noted that the choice to work with children aged 10 to 12 years was particularly fortunate. Besides being at an interesting point in their cognitive development, children of this age are also at an interesting point in their social development. The mentoring provided to these users required less instruction than might have been the case with younger children and in a significant number of instances the sessions felt very much like collaboration between peers.

6.4 Summary

The MachineShop system was tested in an experimental setting where the users interacted with the researcher in a manner similar to the apprentice model of skilled crafts and trades. Six users, 2 girls and 4 boys, aged 10 and 11 years participated in these tests and a total of 8 automata were produced (see Appendix B) over periods of from 13 to 37 weeks. Each user designed at least one automaton and refined that design by iterating over a series of prototypes made primarily from inexpensive materials. Mechanical components were created using the MachineShop software and a computer controlled, carbon dioxide laser cutter was used to fabricate most of the pieces required for construction. Users completed their automata by assembling the parts they had designed and fabricated and, in three instances, by painting and decorating.

Most users showed increases in their mechanical reasoning and spatial cognition abilities, as well as their domain knowledge and their skill in creating automata. There was a marked improvement in the confidence of most users with respect to their abilities in design and construction as well. The results presented here will be illustrated in Chapter 7 where the work of one user is described in more detail.