UNDERSTANDING COGNITIVE OVERLOAD IN GENERATIVE DESIGN

An Epistemic Action Analysis

HALIL ERHAN\textsuperscript{1}, JANELYNN CHAN\textsuperscript{2}, GILBERT FUNG\textsuperscript{3}, NAGHMI SHIREEN\textsuperscript{4} and IVY WANG\textsuperscript{5}
\textsuperscript{1,2,3,4,5}Simon Fraser University, Canada
\textsuperscript{1,2,3,4,5}\{herhan|jlc40|gilbertf|naghmi_shireen|ivy_wang_7\}@sfu.ca

Abstract. Choice overload is experienced when designers use generative systems to explore a large number of alternatives. In an experiment, we studied the epistemic actions designers perform to reduce their cognitive load caused by possible choice overload during design exploration. The participants were asked to select alternatives among a large set of solutions in a simulated design environment. For data encoding, we adapted an epistemic action analysis method to understand which actions occur in what phase of design. Most epistemic actions are observed during criteria applying phase. The most frequent actions were ‘clustering and grouping’ and ‘talking and gestures to guide attention’. Ultimately our goal is to answer if a system can alleviate the possible cognitive overload when working with a large number of alternatives, if so how they would look when implemented.

Keywords. Generative design; parametric modeling; cognitive overload; selection; epistemic actions.

1. Introduction and Motivation

In the presented study, we aim to explore the role of epistemic actions that designers perform to alleviate the cognitive overload that may occur when generating and exploring a large number of design alternatives under a set of (design) choices. An epistemic action’s primary goal is improving cognition by reducing the memory, the number of steps, or the probability of error in mental computation (Kirsh & Maglio 1994). By a ‘design choice’ here, we refer to the salient decisions a designer can possibly make in the course of design that can be observed and evaluated. In parametric design, for example, value assignment to a parameter from an arbitrary range of possible values, deciding to add or subtract a design feature, or selecting a set of derived parameters for creating solutions are design choices.
A design alternative, consequently, is a generated solution using a particular set of design choices. We hypothesize that the possibility of a large number of design alternatives and design choices compounded with the lack of tool support for working with them can cause designers to face with cognitive overload problems. For the purpose of this study we define cognitive overload as the demand made on the cognitive system (Fan & Lei 2006) that exceeds the capacity and the capabilities of the human cognition and the available cognitive artifacts. For example, cognitive overload problem may manifest itself when designers are simultaneously creating and controlling design models that are driven by an arbitrary range of possible parameters, which in turn can result with a rapidly growing solutions.

Lack of tools to support epistemic actions can impede determining the design criteria, creating and evaluating design alternatives, and managing them. The systems for editing a single model (Terry et al. 2004), by description, is not intended for creating multiple alternatives. We contend that the choice overload is expected and normal without a proper and direct system support to navigate through possible choices and alternatives, and to assess their outcome. Finding about the epistemic actions can help us to create design systems where choices can be tested rapidly in large numbers by the tools for supporting such actions. But what are these epistemic actions, when do they appear in the design process, what are the frequencies, and what does trigger them? More importantly, if and how they can be translated into tool features.

As a first step to seek answers to such questions, we conducted an experiment to explore the epistemic actions that may take place when working with a large number of alternatives. The epistemic actions performed in an existing system can be relatively easy to observe. For example, in parametric design, using descriptive names for solutions to group them is a memory-saving action; organizing the layout of a propagation-based graph for reducing later search time of particular parameters is a time-saving action; and testing multiple values for parameters before committing to a value can reduce probability in error in the later stages. However, we have yet to have a system to support working with a large set of alternatives created using a large set of choices. In the study we asked designers to work in a simulated design setting where they are required to identify plausible design solutions among a given large number of possible designs. Along with the basic criteria and parameter set given, they are encouraged to introduce their own criteria if and when needed. We attempted to address the reasons as to why, where, and when cognitive overload problems may occur when working with alternatives, and what actions can reduce such problems.

2. Background: Choice Overload Affecting Decision-making

The choice overload hypothesis has been under evaluation over couple of decades. The hypothesis states that an increase in the number of options to choose from may lead to adverse consequences such as a decrease in the motivation to choose or the satisfaction with the finally chosen option. However, there are contradicting findings in the literature as well. The meta-analysis of the research on this topic by Scheibehenne, Greifeneder & Todd (2010) and later by Chernev, Böckenholt, Goodman (2015) questioned when large choices can improve the output
of the task and when they can create bottlenecks in the process of choice. Iyengar & Lepper (2000) studies the decision complexity caused, at least partially, by the (large) number of available decision alternatives. Most research is conducted in the marketing domains where product choices are presented to the consumers, and possibly not giving enough preparation or assisting tools to manage choosing among the alternatives. They also limit the number of choices to about maximum 30 options so that the decision maker can complete the given task in a strictly defined settings. In design, however we have a different scenario where designers are the ones who define the goals, solutions, and tools to be used in completing the design task, which is usually open-ended and more rewarding due to their creative nature. Hence, the result of choice overload can be a relative phenomena and can be taken differently from one domain to another, and from one task to another.

In data analysis domain, Woods et al. (2002) identified three characteristics of information overload on cognition: (a) there is too much data in the workspace, creating clutter; (b) the abundance of data creates a workload bottleneck because there is too much information to sort through in the available time; (c) finding the significance of the data is a challenge because at the outset it is hard to know where to look. Following can reduce the cognitive overload: (a) reducing the amount of data units displayed; (b) automation or cooperation with other cognitive agents; and (c) representing the data in a way so significant data emerges.

The search to find a solution take place in a problem space (Simon 1955) that can be expanded to include user’s knowledge combined with any available external resources (Ashcraft & Klein 2010) in both the physical and informational states (Kirsh & Maglio 1994). What creates significance for a piece of data in a problem space includes: (a) other related data; (b) how the set of related data can vary with larger context; (c) the goals and expectations of the observer; (d) the state of the problem-solving process and stance of others.

Woods, Patterson & Roth (2002) characterizes three ideas in the nature of data overload faced by our cognitive systems: (a) data overload is difficult because of the context sensitivity problem-meaning lies not in data, but in relationships of data to interests and expectations (b) new waves of technology exacerbate data overload when they ignore or try to finesse context sensitivity; (c) the mechanisms of human perception and cognition that enable people to focus on the relevant subset of the available data despite the fact that what is interesting depends on context. They identified particular constraints that must be met for all potential solutions, particularly the idea of organization preceding selection. They describe a data availability paradox, that as the technology easily increases the data available, the human ability to interpret the data has not kept up. While beneficial in some ways, it is also problematic in how we must have more and more effortful search for finding something meaningful and informative that meets our goals and tasks. Scheibehenne et al. (2002) states that a large assortment of choice make exhaustive search undesirable in terms of time and effort, which could contribute to fears of a less than optimal choice (such as second best). This wide availability of data can be key to our success, but only if we are not made vulnerable by it (Woods et al. 2002).
3. Method

In a simulated task environment, we observed the designers’ behaviour when they were asked to work with a large number of generated design solutions. Note that, the data collected in this study aimed to answer other relevant questions in addition to the ones concerned here. In this study, we mainly used the video recordings, since they directly expose the designers’ moves and their comments. When needed, we turned to the questionnaires and in-depth interviews. We applied a mixed-method approach combining qualitative research analysis with supporting visual analysis and quantitative techniques.

4. Participants and Task

Ten designers volunteered to participate in this study, one of which we used as the pilot case to develop a protocol and to test the interrater reliability. For our purpose and an initial experiment we found this acceptable. We expected to recruit architectural designers with experience in parametric CAD for design exploration. However, due to lack of availability for such participants, we resorted to recruiting for a partial experience match. Four participants had experience in architectural design and parametric CAD. Five participants had experience in other design disciplines (graphic design and industrial design). They all use of various computational tools in their design. One of the participants had experience in real estate and graphics design.

Participants were given 1000+ building design alternatives generated using 27 independent parameters (c.f. Erhan et al. 2014 and Wang 2015). We chose visually diverse alternatives to be as stimulating as possible (figure 1). They were asked to go through the given alternatives and select the possible solutions using their own methods and strategies. They were encouraged to propose additional requirements for the building design. The task was deliberately open to interpretation to ensure that they can make their own choices.

![Sample design alternatives from the set used in the study.](image)

The large number of alternatives and openness ensures the participants are overloaded with choices. In real life, designers may add information to turn more unknowns into knowns. However, more or less parts of the problems always remain to be discovered. Also, designers may be overloaded with a different number of
choices, influenced by the task environment, tools, and of course individual differences. In this study, participants are limited to using alternatives printed on index cards, pen and paper, and whiteboards. Participants were told to expect the average task time to be 1.5 to 2 hours, but they could spend as much or as little time as they wished to. The actual time spent on this task by 10 participants ranged between 0.5 to 3 hours.

5. Procedure
The study was conducted in a small room with two whiteboards on adjacent walls (figure 2-Left). Four cameras were used to record participants’ actions: two recorders (red stars) to capture the two white boards; one (GoPro1) to capture an overview of the table; one (GoPro2) mounted on the forehead to get a general proximity of the participant’s attention. One observer and one note taker were sitting in the corners of the room to be as nonintrusive as possible. The table was angled to give ease of access to both white boards. Figure 2-Right shows time-lapsed view from the overhead camera, with five minutes of footage overlaid on a single image frame.

6. Analysis Methods and Process
We adapted the epistemic actions analysis developed by Esteves et al. (2015) that focuses on a set of epistemic actions identified through an extensive literature review, and measures different epistemic actions during problem-solving tasks. While the method designed for ‘tangible interaction’, its generality is based on problem solving and hence is suitable also for analyzing the design tasks. Out of the proposed 20 epistemic actions, we selected 18 after our first pass on the videos (Table 1). Previously, Wang (2015) identified four different phases by analyzing the same videos. She found that the participants commonly follow the same cycle consistent with design literature for describing design exploration (Maher & Poon 1996; Brown 2009). Evidence from her study confirms that designing with generated alternatives follows the same cyclic pattern as general design exploration process (figure 3). Below are the definitions of the phases (Wang 2015).
Figure 3. Observed cycle for design exploration with a large number of alternatives.

Table 1. The epistemic actions and possible interpretation.

<table>
<thead>
<tr>
<th>Epistemic Action</th>
<th>Example interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Clustering or grouping artifacts together</td>
<td>Adding a car or cards to pile</td>
</tr>
<tr>
<td>2 Spatial arrangement of artifacts in relation to one another, the task environment or the user(s)</td>
<td>Creating piles of cards</td>
</tr>
<tr>
<td>3 Rearrange a representation</td>
<td>Shift piles or cards around</td>
</tr>
<tr>
<td>4 Manipulation of an artifact</td>
<td>Moving, rotating a card, taking or dividing handfuls from a stack</td>
</tr>
<tr>
<td>5 Tag or annotate an artifact</td>
<td>Write on a card</td>
</tr>
<tr>
<td>6 Using the body to externalize an internal process</td>
<td>Talking to oneself</td>
</tr>
<tr>
<td>7 Annotate</td>
<td>Write on something not on a card, e.g. sticky, paper, whiteboard</td>
</tr>
<tr>
<td>8 Bodily marking an artifact</td>
<td>Keeping finger on a card, keeping a card close to body</td>
</tr>
<tr>
<td>9 Build a model or external representation</td>
<td>Explicitly establish a criterion</td>
</tr>
<tr>
<td>10 Bimanual use of two artifacts, two representations, or an artifact and a representation</td>
<td>Holding or comparing two cards in hand</td>
</tr>
<tr>
<td>11 Divide workspace into several stations in which only a subset of actions is afforded</td>
<td>Interpreted same as 12</td>
</tr>
<tr>
<td>12 Place artifact in a contrasting environment</td>
<td>Moving piles or cards into a separate distinct area e.g. move from table to whiteboard</td>
</tr>
<tr>
<td>13 Compare an artifact with a possible destination or other artifacts</td>
<td>Interpreted same as 10</td>
</tr>
<tr>
<td>14 Talking or gesturing to guide and direct attention</td>
<td>Talking or gesturing to researchers</td>
</tr>
<tr>
<td>15 Use of an artifact or tool to physically constraint the user, or the user of other artifacts and tools</td>
<td>Using tools to perform an action or help decision making, e.g. use a ruler to estimate scale</td>
</tr>
<tr>
<td>16 Clear and clean clutter</td>
<td>Remove discard piles or large quantities of cards from workspace</td>
</tr>
<tr>
<td>17 Shuffle artifacts</td>
<td>By description</td>
</tr>
<tr>
<td>18 Test the state or response of a system, model</td>
<td>Review or check brief sheet</td>
</tr>
</tbody>
</table>

**Criteria (re)Building**: Define a criterion to decide how to treat a given set of alternatives including when a criterion is replaced with a different criterion, rearranging criteria, eliminating a criterion or redefining its meanings.

**Criteria (re)Applying**: Reconsider the qualities of an alternative against one of the chosen criterion.

**Criteria Testing**: After introducing a criterion, test the first round of selections’ acceptability. We observed multiple instances where a criterion tested and
Reflection and Resetting: Take pauses to reflect on the process to explain what have done so far, what would be the next move. The designers may continue or may simply reject the previous moves and restart with a new strategy.

Two encoders separately analyzed the pilot case following the initial protocol established. The interrater reliability (89% Cronbach’s Alpha) was achieved in 10 minute intervals by looking for 18 epistemic actions in the four predefined categories. The actions were identified by both visual descriptions recorded (Table 1) as well as the verbal descriptions of the participants at the time when the action is observed. The pilot study revealed that not every action had a corresponding category, since some actions failed to meet the predefined categories. We added two new categories: (a) task evaluation, which refers to the actions that defines the objectives of the new task, and typically prior to building any particular criteria; and (b) task organization that includes actions observed when the participants already set their objective but before performing actions outside of criteria building and applying.

The data recorded includes the time when the action is observed, the duration of the action, the category of the phase in which the action is observed, and the researchers note on that particular action. We analyzed the data after combining the findings from each encoder and normalizing the task time.

7. Findings

We observed that most frequent epistemic actions are ‘clustering and grouping’, ‘using body to externalize an internal process’, and ‘talking or gesture to guide attention’ (figure 4). Although, higher frequency does not necessarily mean that these are the most decisive moves, they definitely help managing the cognitive load. The last two reveals ‘thinking’ about the actions taken or to be taken. The least observed actions were ‘placing artifact in a contrasting environment’ and ‘building a model or external representation’. ‘Tagging and annotating artifacts’ and ‘dividing workspace into stations’ are rare actions.

Criteria Building is the most dominant phase at the start of the task, which gradually decreases (figure 5). Six participants built criteria at or near the beginning and once again towards the end (figure 6). Building of criteria at the start could be interpreted as formulation of a model for the selection process to follow. Criteria building at the end could be interpreted as formulation or revisiting of discarded items as “bad examples”, which is the analysis of non-candidate solutions as a strategy to manage data overload. The participants may have employed this strategy towards the end as a means to lighten their cognitive load and organize their thoughts for explanation.

Reflection is also seen throughout, though most commonly seems to occur after spurts of Criteria Applying. This may be a self-evaluation or evaluation of the just performed action in how much or how in line it is to expectations. Task Evaluation is expectedly at the start of the task, when participants are familiarizing themselves with the expectations, environment, task etc. Task Organization is quite spread out, some notably occurring after criteria applying. This may be due to simple
neatening and straightening up of item stacks or may be reorientations or refreshing of perspective, another effective constraint against data overload as pointed out by Woods et al. (2002).

In **Criteria Applying**, the most frequent actions are ‘clustering or grouping artifacts together’ and ‘using body to externalize an internal process’. In ‘Reflection’ phase, however, particularly ‘talking and gesturing’ appear more than the others. This is consistent with the literature as it requires less effort than other actions in reflecting in the process. We wonder, however, if the participants were asked to describe their reflection on their actions through other means but talking, how would they explain their choices.

Eight participants reported having difficulty narrowing down alternatives. Some participants referred to the task as more focused on grouping rather than elimination. Initial criteria development was challenging due to lack of constraints and knowledge about the generated sets. Participants made their own assumptions of the situation and define their own rules and understanding of the set. There are difficulties in deciding about each individual alternative. Participants reported “sitting on the fence” for many of the alternatives and “questioning about the se-
lection, [...] if it’s an error, yea, definitely an error”. Also as a concept matured, adding or removing alternatives from a group of selected criteria becomes harder due to the fear of having too many elements confusing their intent: “I want to make my elements of design as clean as possible, not too much junk, [...] my interest is in this kind of geometry, now is the new addition going to help to clarify it, or is it going to be two different kind of designs?” Three participants mentioned a desire to go through all 1000 alternatives to avoid “missing out on something interesting.” One participant that finished viewing all 1000 alternatives in the session.

Figure 6. Epistemic actions performed during each predefined phase (use figure 5 as legend).

8. Conclusions and Discussion
We examined the epistemic actions designers perform to manage their choice overload that may be experienced when working with a large number of alternatives. We borrowed a list of epistemic actions from Esteves et al. (2015) and design phases from Wang (2015) in encoding the data. We identified the epistemic actions, when they take place, and how often the designers rely on them. The findings show that criteria applying is the most intense phase in design with possibly clustering and grouping is the most frequent action observed. These actions should be treated differently in designing a new system. While participants were encouraged
to describe their actions, the encoders might confused them with the ‘talk and reflect’ epistemic action. Identifying its effect on the action count and categorization would be difficult; however, their talk (aloud) might have helped unintentionally in making the selection of alternatives. This may be a discussion point in further studies.

While computational tools have potential to generate a large space of solutions, they provide almost no direct features for the epistemic actions in the natural workflow of design. We can integrate design generation and exploration by eliminating the extra task layers imposed by most systems. The system features for this can be revealed by observing how designers act when they make design choices. To alleviate the hypothesized cognitive overload problem, what are the functional and structural system properties? How can we identify these properties and how could they be implemented? We leave the discussion on these questions to another paper.

References