Stalking homo faber: 
A comparison of research strategies for studying design behavior

David Latch Craig 
College of Architecture 
Georgia Institute of Technology, Atlanta, USA

Abstract
Four different strategies for studying design behavior are examined and compared: think-aloud protocols, content analysis, process isolation and situated studies. The four strategies span a range of data collection and analysis methods. More substantial differences, however, are found to exist in the theoretical biases each entails. Think-aloud protocols assume that design can be described as search; content analysis assumes that the structure and content of internal representations is critical to understanding problem-solving behavior; process isolation assumes that design can be cognitively decomposed; and, finally, situated studies assume that a better understanding of design involves the study of designers and their environments as integrated systems. Despite their sometimes competing claims, these four strategies can potentially work together in complementary ways to provide insights into design behavior.

1. Introduction
How do designers solve design problems? What exactly is it that designers do? Researchers in a variety of fields have sought answers to these questions for a variety of different reasons. Researchers in cognitive psychology, for example, have looked at design both because it represents a core aspect of human creativity and because of the challenge it poses in applying what is already known about cognition to more complex domains. Researchers in fields like engineering, computer science and architecture, on the other hand, have looked at design because of a larger interest in improving what designers do through the use of design aids and better teaching methods. At the same time, design has been the subject of study in a host of other fields, by researchers pursuing their own discipline-specific goals, ranging from the sociological to the anthropological to the historical.

The discussion of research strategies in this chapter is not limited by disciplinary boundaries. It is, however, limited by the general goals of research. In particular, the chapter addresses empirical research strategies aimed at understanding how designers solve design problems. Despite this specific focus, a lot of room is left for diverse strategies to be played out, particularly given the scope and complexity of the subject
under study. Design problems are not solved in any simple or unitary way. They are solved, rather, in a range of social, cultural and physical settings, through the efforts of individuals as well as groups, using a variety of different types of knowledge and skills.

In all, four research strategies are considered: think-aloud protocols, content analysis, process isolation and situated studies. Although a variety of published studies are discussed using this classification scheme, not all fit neatly within it. Most studies make use of several strategies at once, resulting in almost as many distinct approaches as there are studies themselves. Moreover, this classification scheme does not necessarily represent different data collection techniques or methods of analysis. The strategies differ, rather, primarily in terms of their theoretical orientations. Although some studies appear schizophrenic in their use of theory and method, theoretical orientation is noted wherever it is made explicit in published reports. In some cases, liberty has been taken in reclassifying studies to help illustrate less conspicuous research alternatives. Along these lines, one of the goals of the chapter is to help strengthen strategies that have been used by researchers but, for lack of theoretical and methodological clarity, may not have been used to their full potential.

Before getting into the research strategies themselves, two assumptions will be made, first that design is a distinct type of behavior and second, that design behavior is open to description. Although the first assumption is subject to debate, no special claims will be made about what design is at this point. Design will simply be considered any problem-solving activity that results in the creation of an artifact or a plan for generating an artifact. Different research strategies, of course, have different things to say about what design is, as will become evident later in the chapter. The second assumption is somewhat less controversial. Although at times it may seem that design follows divine inspiration, that it is ineffable or that it withers under description, there are too many levels at which one can explain human behavior and too many techniques for elucidating reasoning to make any claims against empirical research outright. Not all research strategies, of course, will be effective, nor should we assume that a single all-encompassing description of design is realizable. Some strategies, however, may be more appropriate for studying design than others, and the purpose of the chapter is precisely to see how different strategies stack up.

2. Think-aloud protocols

In what is perhaps the most thorough account of how one might study problem-solving behavior, Ericsson and Simon (84) present a research strategy that hinges on having subjects talk aloud or think aloud while executing a task. The goal of the strategy is to track and better understand the sorts of reasoning processes and strategies employed by problem solvers in general. The “think-aloud” version assumes that some mapping may be required between verbalizations and internal representations, whereas the “talk-aloud” version assumes that internal representations are already in verbal form. “Thinking aloud” is thus more general and will be used here to refer to both methods. Although other strategies for studying cognition and behavior require subjects to talk while performing a task, think-aloud protocols will be used to refer specifically to methods that use concurrent verbalization to track thought processes.
Ericsson and Simon’s defense of think-aloud protocols begins with the observation that all data collected in behavioral experiments, whether verbal or not, is subject to interpretation. Hence, there is no a priori reason to assume that data collected in think-aloud protocols will be any less objective than data collected in other experiments. Ericsson and Simon do not assume that subjects have direct access to their mental processes, nor do they assume that thought is “inner speech.” They only assume, consistent with their view of verbal reports as behavior, that one can infer upon hearing a subject say “X” that the subject was able to say “X.” Following accepted cognitive theory, one can then infer that if the subject was able to say “X,” the subject had an informationally equivalent representation of “X” active in short-term memory (STM). And, finally, if one also accepts a serial information-processing view of cognition (i.e., that thought consists of “a sequence of internal states successively transformed by a series of information processes”), one can infer that “X” is somehow related to what was stated earlier and what will come later.

Ultimately, think-aloud protocols require the interpretation of data at two somewhat distinct levels of theory in order to infer underlying thought processes. The first level is low-level cognitive theory of the sort mentioned above, while the second level is the level at which hypotheses are made. The two may interact but in a potentially positive way. Ericsson and Simon refer to this as “bootstrapping”: “We need a model in order to interpret data that are to be used, in turn, to test the model. Under these circumstances, our data-interpretation model should be as simple as possible, and it must not incorporate components that are themselves bones of theoretical contention” (Ericsson and Simon 84, p. 10).

The main difficulty in making use of think-aloud protocols lies with the basic task of interpreting utterances and relating them to prior utterances. This is where reliance on higher-level theories comes into play. According to Ericsson and Simon, protocol data should be coded based on a theory of what sorts of things a subject might reasonably be thinking about. In the wealth of protocol studies conducted by Newell and Simon (1972), for example, this entails mapping out plausible problem spaces that subjects might traverse in the course of completing a task. Plausible problem spaces, in turn, yield a list of concepts that will be relevant, which then serves as a “dictionary” for interpreting utterances. The a priori formalization of an encoding will, of course, help guide an interpretation even if the definitions are open-ended. For example, Ericsson and Simon point out that a particular mapping from linguistic utterances to concepts might result in identical translations for “red circles are zog” and “blood-colored circles are zog.” “Blood-colored” can be assumed to contain the same information as “red” as long as one’s high-level theory of what the subject is thinking does not count the two as being different.

Think-aloud protocols appear to work reasonably well when studying how subjects solve well-defined problems like the Tower of Hanoi and cryptarithmetic puzzles (Newell and Simon 72). In such cases, the way the problem is presented to subjects leaves little freedom in representing the space of possible moves and solutions. Thus, “whenever verbalizations correspond to plausible intermediate states in a processing model for the problem-solving activity, we can plausibly infer that this information is
actually used in generating the problem solution” (Ericsson and Simon 84, p. 171).
From there, one can generate hypotheses about how a subject might navigate or 
transform a problem space (e.g., what sort of heuristics will be used) and compare these 
with how a subject actually traverses or transforms such a space.

One criticism of think-aloud protocols is that although verbal data can be used for a 
variety of purposes, it may be inappropriate for tracking sequential thought processes. In 
particular, think-aloud protocols may misrepresent underlying processes to the extent 
that subjects “sometimes cannot report on the existence of critical stimuli, sometimes 
cannot report on the existence of their responses, and sometimes cannot even report that 
an inferential process of any kind has occurred” (Nisbett and Wilson 77, p. 233). 
Ericsson and Simon respond to such attacks by arguing that it is only in experiments 
where subjects are explicitly asked to explain why they formed certain judgments or 
behaved in a certain way that fallacious reports are generated. This, they point out, is not 
consistent with their proposed use of verbal data. They never claim that subjects have 
access to cognitive processes; at best, they claim subject can report the contents of STM. 
In Nisbett and Wilson’s study, all accounts are retrospective, in which case it is possible 
that traces of earlier processing have disappeared and must therefore be inferred. 
Ericsson and Simon specifically note the potential dangers of retrospection: Subjects 
may retrieve the wrong episode, they may retrieve information that is related to but not 
part of a prior episode, they may simply infer a plausible course of reasoning, or, finally, 
they may unconsciously neaten up an episode, leaving out missteps and reasoning errors 
that were actually part of their reasoning process.

Other criticisms of think-aloud protocols are that verbalization provides an incomplete 
account of the reasoning undertaken by subjects even if verbal reports are accurate, and 
that verbalization may potentially interfere with thinking processes in other modalities. 
Both of these criticisms are limited if one is strictly dealing with well-defined tasks. If, 
in other words, a subject is given enough information to clearly specify a problem space, 
gaps in the protocol may be interpolated since, as evidenced by a task analysis, only a 
limited set of relevant information states can be instantiated. If a subject’s thoughts 
wlter to last night’s baseball scores while solving a well-defined puzzle, we can rest 
assured that those thoughts will not affect the puzzle-solving process in any qualitative 
way. In the same sense, verbalization is less likely to interact with reasoning in other 
modalities if one is applying stepwise procedures towards achieving a goal (Schooler 
and Melcher 95). Verbalization, if anything, may help subjects reflect on the state they 
are in and develop a more explicit map of a well-defined space.

The incompleteness of think-aloud protocols may be more problematic, however, if the 
problem space is not well defined. In such cases, the experimenter must simultaneously 
infer both the problem space and the state the subject is in from the verbal data. Ericsson 
and Simon recognize this problem, noting that in actuality problem spaces associated 
with even the most well-defined tasks can change unpredictably over time. When faced 
with an unclear problem space, Ericsson and Simon suggest using the weaker criteria of 
relevance and sequential consistency to establish the contents of and links between 
information states. They add, however, that these criteria preclude the study of “relatively unconstrained cognitive activities like free association or daydreaming”
Knowing and Learning to Design

(Ericsson and Simon 84, p. 171), as well as processes that are not cumulative. In this case, despite Ericsson and Simon’s general attempt to rescue think-aloud protocols from problem space uncertainty, problems may still persist in coding design protocols. For one thing, design may not involve strictly cumulative cognitive processes. A designer’s thoughts may wander laterally, and free-association and daydreaming may, in fact, be useful. This, of course, amplifies the problem of inferring what a designer is thinking and how it fits in with a longer train of thought. One might rely strictly on the local context of utterances to infer the knowledge state of a designer, as suggested by Dorst and Dijkhuis (1996), but such an approach seems to trivialize the nature of internal representations. Internal representations are potentially complex, so much so that verbalization may only reveal the tip of the representational iceberg (see, e.g. Barsalou 93).

In addition to incompleteness, interference with other modalities may also create problems for think-aloud protocols in ill-defined domains or, for that matter, any domain where subjects are not able to make use of strategies that correlate well with speech. Think-aloud instructions in insight problems have, in fact, been shown to hamper problem-solving ability, beyond simply slowing down the completion of a task (Schooler and Melcher 95). Although it is not clear what this says about the underlying reasoning processes, verbalization may tie up resources that could be used in other ways, or it may bias particular (e.g., logical) reasoning strategies over others. Among other things, experiments have shown that verbal descriptions of visual imagery impair memory for such scenes (Schooler and Engstler-Schooler 90; but see Freund, cited in Loftus and Bell, 75, for different results). Hence one might assume that, at a minimum, verbalization will limit a problem solver’s ability to take full advantage of visual information.

The next section reviews the use of think-aloud protocols in design and addresses their potential limitations in more detail. Although, as mentioned earlier, there are a variety of research strategies that require subjects to talk while performing a task, the studies referenced here all make use of Ericsson and Simon’s proposed strategy for using think-aloud protocols.

2.1 Think-aloud protocols in design

Think-aloud protocols have been employed in design studies for more than thirty years. In that time, a variety of distinct approaches have been taken. All of the studies referred to in this section describe design behavior as movement within a problem space and use protocol data either to illustrate, support or expand cognitive models of design. To help organize the discussion, the studies are grouped into two categories: those that describe search in terms of a set of operators or production rules for generating moves, and those that treat search more generally, using quantitative analysis rather than model instantiation to identify patterns of behavior.

2.1.1 Production-rule approaches

The earliest application of think-aloud protocols in design (Eastman 69) and some later applications (e.g., Akin 78; Ullman et al. 88) model the design process with a set of
cognitive operators that generate and transform information about the problem or about a developing solution. The problem space, in this case, is assumed to be locally rather than globally defined, consistent with Simon's (1973) description of ill-structured problem-solving processes. In other words, information is assumed to be attended to and manipulated in a limited fashion, such that the problem space changes continuously throughout a problem-solving episode as new information comes into view. Operators that are postulated typically deal with things like the retrieval of requirements, constraints and general knowledge from long-term memory (LTM); the assimilation of information from external sources (e.g., books); the transformation of such information through deductive inference or through the recognition of patterns; the transformation of requirements into form; and finally, the fine tuning of form variables to match constraints. In simpler terms, production-rule models argue that constraints and requirements are transformed into information that can then be used to specify forms. The latter usually lead to multiple alternatives, since form variables and subtype specifications are often left open. Using an example from Akin (1978), the requirement that sleep be supported in the design of a house can be (partially) transformed into the specification of a bed, which can then be varied in terms of size, style and placement.

Design processes modeled on production rules are not necessarily deterministic. Because the space of potential design decisions can be large and because designers may not have enough expertise in a domain to follow a set decomposition, the result can be idiosyncratic. The kinds of information initially brought into the problem space and the sequence in which requirements and constraints are addressed may vary, depending on personal experiences, on the structure of information in LTM and on the retrieval cues present in the environment. As Ullman et al. (1988) points out, a design process that results in a surprising or innovative design can be modeled with production rules to the extent that the decomposition need not be known in advance. Aside from this variability, however, design processes modeled using production rules are analogous to those associated with solving puzzles: Some complex combination of component types and variables must ultimately be specified, usually incrementally, to satisfy a set of requirements and constraints. Although rules can also be used in design to transform requirements and constraints themselves (which is not usually the case with simple puzzles), this can be seen as simply adding an additional layer of complexity to the problem-solving process.

Although think-aloud protocols have been used to illustrate production-rule models and in some cases to expose novel search strategies (e.g., Akin 78), there are several questions concerning this strategy. In particular, one might ask whether rule-based knowledge is sufficient for describing the sorts of reasoning that go on in design and, hence, whether think-aloud protocols coded using production-rule models accurately or fully capture design reasoning. Although rule-based reasoning is superficially evident in most protocols (in the form of rationale like, "I'll add a bed here because..."), this sort of reasoning may simply mask the use of more complex relational knowledge structures (Barsalou 93). Notably, it has been found that when individuals are put in controlled situations that trigger the retrieval of specific knowledge structures, or schemas, without their knowing, they often manufacture rationale after the fact to explain subsequent judgments simply because they have no other way to account for their reasoning.
Knowing and Learning to Design

(Nisbett and Ross 80). Their rationale, moreover, may not be consistent with their actions. Although, as mentioned earlier, such research relies on error-prone retrospective accounts, the issue here is not simply whether verbal reports are correct or incorrect. The point, rather, is that reasoning may be buried in extended knowledge structures that do not surface in the form of isolable, easily verbalized rules.

The primary difference between reasoning from schemas and reasoning from rules is that schemas carry with them a halo of unconscious, but often consequential, assumptions. If one reasons about a particular person, for example, one might reason from a stereotype and actually infer features or traits that are not present. In design, the complexity and variety of retrieved and assimilated information suggests that a lot of unconscious assumptions are folded into the problem-solving process. Some unconscious assumptions amount to little more than the instantiation of default values in knowledge structures. For example, one might add a bed to a room and assume that it is white without consciously considering color as a variable. Designers might also, however, invoke knowledge structures from analogous domains, in which case conceptual constraints and structural relations may direct the development of a solution unconsciously. In Ullman et al. (1988), for example, a designer working on the design of a mechanical device assumes early on that one component will be used to satisfy two crucial functions, only to overcome this self-imposed restriction late in the design process. The subject, in this case, appears to have retrieved and utilized a schema for a particular kind of mechanism without actually considering the structure of the retrieved representation or the fact that a specific schema was retrieved at all. The representation, in other words, guided the subject’s behavior without the subject consciously making any decisions.

The use of think-aloud protocols for illustrating, testing and exploring production-rule models of design comes down to the ability of the experimenter to infer a potentially large and complex set of connections between utterances made by subjects. This requires empathy, since the problem space is not made explicit in the problem statement but rather develops out of the designer’s actions. As Eastman points out, “the analyst can understand the problem solver’s processes to the degree that he can find correspondences between the processes he has experienced and thus understands, and those of the subject” (Eastman 69, p. 670). The problem, in this case, is that even if the experimenter is empathetic, he or she may simply be making the same unconscious assumptions that the subject makes and hence fail to uncover all of the underlying processes. Given this limitation, it is tempting to describe the use of think-aloud protocols to model processes in design as "concurrent introspection." There are, of course, design domains in which expert knowledge is well codified, and there are also design problems for which components are completely (or largely) given in advance and for which variables and subtype specifications are limited. Both would potentially make the task of inferring thought processes from think-aloud protocols more plausible. This, for example, may be the case in Eastman’s (1969) study of the design of bathrooms given a fixed budget and a fixed enclosure. However, even in the case of highly constrained bathrooms, everyday concepts about bathroom use, which may carry a host of unanalyzed assumptions, can enter into the process.
2.1.2 Quantitative approaches

Other design studies that utilize think-aloud protocols of the sort described in Ericsson and Simon (1984) have taken a much looser approach to analyzing behavior. In particular, some studies have taken a quantitative approach, where utterances are segmented, categorized and analyzed for correlations, frequencies and patterns (e.g., Akin and Lin 96; Gero and McNeill 98; Goel 95; Goldschmidt 96, 97; Lloyd and Scott 94). Such studies rarely concentrate on complete cognitive models of design. Instead, they begin with intuitive ideas about things like the fomentation of critical decisions or the behavioral underpinnings of effective problem solving, and then use pattern analysis to fill in the details. Although starting off without well-defined hypotheses may suggest that this strategy is vulnerable to confirmation bias, in practice it is analogous to an archeologist combing through layer upon layer of sedimentation. Hypotheses that are specific to the data must take shape as the data becomes available simply because data for answering specific questions cannot be generated willfully. In this spirit, quantitative approaches are less likely to be called experiments and more likely to be called "macroanalysis" or "measuring designing" (Gero and McNeill 98).

Although the quantitative studies discussed in this section focus on the structure of the design process, they nonetheless employ search space models of problem solving at the outset and frequently refer to Ericsson and Simon (1984) for methodological support. They differ from the studies discussed in the previous section, though, in that they look for statistical correlations and patterns within verbal reports, sometimes on the assumption that the design process is too unstructured to be tracked using production-rule models. In this sense, they acknowledge that subjective and often unconscious knowledge structures wreak havoc on detailed analyses and thus drop content altogether. In the case of Goldschmidt (1996), for example, every coherent statement made by the designer is assumed to constitute a move that can be linked to other moves in unpredictable (both to the designer and to the experimenter) ways. Links are not coded as a sequence of moves simply because it is considered too difficult to say where they came from.

The problem with maintaining a problem space model of design while focusing on statistical correlations and patterns within utterances is that it can be difficult to tell exactly what a move, information state or operator is. In traditional problem space studies (e.g., Newell and Simon 72), a state in a problem space is pinpointed by actually tracing a sequence of moves. In chess, for example, potential states are those that can exist in a legal game, certain configurations of chess pieces simply being impossible. By taking a quantitative approach, however, all evidence of a tractable problem space is blurred, making the assessment of a given utterance especially difficult. In this case, encoding a designer's utterances as evidence of movement through an unknown problem space seems unfortunately like trying to follow the movement of a mouse through an invisible maze to better understand the mouse’s behavior: If the maze cannot be seen, every twitch could be read as a purposeful move and vice versa.

A second concern with the quantitative analysis of think-aloud protocols is that frequencies, correlations and patterns do always help to clarify causal links. Aside from the intuitive theories that accompany such studies, little is ventured about the cognitive
processes or mental representations that might underlie reasoning in design. Patterns or
correlations found in protocol data could simply reflect the conceptual bias of the coding
scheme or, at best, indicate a relationship between behaviors without suggesting how
they are related. Quantitative analysis is generally useful when comparing two refutable
hypotheses. This, however, is not usually the case in design studies. Most hypotheses
deal indirectly with the data. Goldschmidt (1996), for example, postulates that “effective
design processes are characterized by a high ratio of interlinkings among its moves” (p.
74). Her study, however, does not look at effectiveness but rather at how "link densities"
and other patterns compare between individuals and groups. As another example, Akin
and Lin (1996) speculate on the significance of multi-modal behavior (e.g., drawing,
examining and thinking at the same time) and find that it is more likely to coincide with
"novel design decisions" than with routine decisions. They concede, though, that their
analysis does not suggest which fosters the other or if, in fact, they are causally related
at all.

A final concern with quantitative analyses involves the nature of the coding schemes. As
mentioned above, coding schemes are created to partition utterances into categories of
behavior, usually at a level of abstraction above production-rule models. Once a coding
scheme is created, agreement between independent coders is often measured to see how
sharp the scheme is. One problem is that lack of agreement – as low as 60% in some
studies (see, e.g., Lloyd and Scott 94) – does not always deter researchers from using a
particular scheme. A more insidious problem, though, is that the definition of a coding
scheme may be conceptually tenuous, confounding findings produced by it. As one
example, Goel (1995) uses a quantitative analysis to comparatively study movement
between design “modules” (which he defines as distinct design components or design
issues) in well-defined and ill-defined tasks. The question, in this case, is what
constitutes a module for a given task. Goel makes the assumption that in cryptarithmetic
problems (which require subjects to decode letters into numbers in order to satisfy an
arithmetic expression, such as SEND+MORE=MONEY), the columns of the arithmetic
expression, as opposed to the individual letters or letter types, are modules. This coding
assumption directly impacts Goel’s findings, which deal with the frequency of moves
between modules. Although the assumption is contestable, it gets little theoretical
support in the study. Following Ericsson and Simon’s (1984) discussion of the think-
aloud protocols, utilizing theoretically unclear coding schemes may lead to
“bootstrapping” failures and ultimately jeopardize the results of such studies.

3. Content analysis
The notion of relational knowledge structures, or schemas, that Nisbett and Ross (1980)
invoke to explain implicit inferential behavior has developed over the past several
decades into a fairly detailed theory of conceptual structure and conceptual reasoning.
The basic idea is that concepts get their meaning from links with other concepts, often
inheriting or sharing whole complexes of links with other domains, near and far. Links
may be recursive (e.g., humans have progenitors which are also humans) and may be
constrained in complex ways (e.g., a ski vacation has a location which must be snowy).
Some knowledge researchers have worked to develop detailed computational structures
that are capable of accounting for a wide variety of reasoning in a unified way (e.g.,
Minsky 75; Rumelhart and Ortony 77; Schank and Abelson 77; Barsalou 93), while others have used relational structures to analyze new experimental data, for example in conceptual combination tasks (Wisneiwski 97). Likewise, in design, schema retrieval, combination and modification has been used to describe how designers develop solutions to problems in ill-defined domains (e.g., Gero 90).

A research strategy designed specifically to highlight conceptual structure and conceptual change, as opposed to rule-based processes, is what Chi (1997) refers to as “verbal analysis.” While Chi’s interest is primarily in education – that is, in how students’ representations change over time, how they vary in different contexts and how they differ from those of an expert – her strategy for studying problem solving can also be applied to design. In this section, verbal analysis is used as a primary example of a more general strategy that will be referred to as content analysis. Unlike think-aloud protocols, content analysis does not aim to track cognitive processes, although it may still rely on verbal reports for data. Whereas Newell and Simon (1972), for example, perform a task analysis to deduce probable problem space representations prior to a problem-solving episode, content analysis uses verbal data and other externalizations like sketches to infer representations after the fact. This is argued to be particularly relevant in ill-defined domains where problem spaces depend on more than that which is given in a problem statement. In the end, Chi argues that processes are secondary, since “it is the solver’s representation that determines the problem-solving process” (p. 278).

To the extent that content analysis aims to uncover the structure and content of representations, it relies heavily on the qualitative analysis of external, often verbal, reports. However, as Chi points out, quantitative analysis may be used to back qualitative findings. More precisely, “one tabulates, counts and draws relations between the occurrences of different kinds of utterances to reduce the subjectiveness of qualitative coding” (p. 273). An example of this might be counting how many times a problem solver refers to prior problems or examples and comparing those results with likelihood of success on the current problem. Counting, in such cases, usually involves predicting some behavior based on what a subject is thought to know. Such knowledge can be also be tracked by cataloguing the content of utterances and the conceptual links between them. Important questions might, for example, be whether the conceptual structure boils down to a list of perceptual features, a set of fragmented beliefs, a set of central exemplars, or an analogous concept. Each of these may have different consequences for the type of reasoning that follows and ultimately for the way an individual solves a problem. A designer, for example, who works with a set of fragmented beliefs may utilize information erratically or inefficiently, whereas a designer who works from a well-understood analogous concept may be able to manage a variety of concerns and insights by integrating them into an already coherent conceptual structure.

Like think-aloud protocols, content analysis may involve asking subjects to say what they are thinking while solving a problem. However, since the aim of content analysis is to uncover internal representations, some probing on the part of the experimenter is usually admitted. Obviously, probing may affect performance and may alter the underlying processes. This, however, is seen as consistent with the overall goals of
content analysis for two reasons. First, processes are not the primary focus, and thus may be sacrificed if necessary. Asking a designer specific questions about a problem or about existing concepts without actually having them carry out a design might still reveal what sort of knowledge they would use. Second, interfering with underlying processes as they occur may be beneficial provided that the interference can be tracked. One goal of content analysis is to understand how representations change through reflection; hence instigating change may lead to valuable insights.

Also like think-aloud protocols, content analysis relies on theory for the coding of external reports. Although Chi criticizes think-aloud protocol studies for their lack of a control condition (as Chi points out, “because one can always tweak the model to ‘match’ the protocol, it’s never clear how the model can ever be invalidated”), think-aloud protocols are not necessarily aimed at picking out hypotheses as much as they are at organizing behavior in a way that makes discovering higher-level patterns of behavior easier. Content analysis is similarly dependent on theory for telling us what to pay attention to and what to encode as part of a representation. Even if quantitative performance measures are incorporated into content analysis, some faith is required to accept that the variables being measured are sufficiently independent and significant enough to warrant studying. Thus, content analysis is not necessarily a more rigorous research strategy. Content analysis and think-aloud protocols are perhaps best considered complementary strategies, the former being tailored for ill-defined domains and investigating the formation and use of knowledge representations, the latter being tailored for well-defined domains and investigating step-wise processes.

3.1 Content analysis in design

To a certain extent, the quantitative analysis of think-aloud protocols discussed in Section 2.1.2 resembles content analysis, particularly verbal analysis. Both develop quantitative coding schemes based on preliminary qualitative analyses and work between the two reflexively. They differ, however, in that think-aloud protocols focus almost exclusively on processes at the expense of representational structure, while verbal analysis does the opposite. Content analysis, moreover, encourages the use of probes to elicit particular kinds of externalizations.

In practice, few design researchers have undertaken studies that delve into the structure and use of knowledge representations through verbal reports or other externalizations. In part, this may be because designers are assumed to be generalists, lacking any easily pinpointed expertise. Chi, in her own studies, has used verbal analysis in relatively limited domains like physics (e.g., Chi et al. 81). In such cases, even though subjects may lack formalized concepts, there is always an ideal, or target knowledge base against which they can be compared. Content analysis in design, by contrast, must deal with the sorts of elusive references and vague or open-ended concepts that tend to play a role in design problem-solving.

One study that does, however, attempt to uncover knowledge representations used by designers was done by Akin (1978). In one experiment reported in that study, the chunking of architectural plans by architects was investigated, following Chase and
Simon’s (1973) analysis of chess board configurations by chess players. Specifically, architects were asked to look at plan drawings and either trace, copy or interpret them, all while verbally reporting their actions. Akin analyzed each subject’s behavior by noting pauses in their protocols, ultimately inferring the low-level organization of plan knowledge that they used. In a second experiment, Akin asked architects to try to imagine the contents of a drawing by asking questions about it but without actually seeing it. By analyzing the content and sequence of questions that each subject asked, Akin inferred the structure and content of the inferential knowledge they used.

Another study that focuses on the knowledge used by designers more than on actual design processes is reported in Crismond (1997). In that study, a variety of novice and expert designers were given “investigate and redesign” tasks. Initially, subjects were presented with three unlabeled artifacts and asked to identify them. Although the artifacts all performed the same function (opening jars in one series and cracking nuts in another), the subjects knew nothing about them in advance. In addition to being asked what the artifacts did, subjects were also asked a variety of probing questions such as how much they thought the artifacts cost and what type of consumers they were targeted towards. Subjects were subsequently asked to rank the sample artifacts and redesign them. By analyzing verbal reports, Crismond was able to characterize the kinds of device knowledge, scientific knowledge and inquiry skills his subjects possessed. He was also able to track the mapping of knowledge from related domains, finding in one case that experts were more likely to have high-level schemas onto which they could attach and hence integrate individual insights generated from exploratory activities.

Another study, done by Atman et al. (1999), focuses on, among other things, activities that help bring additional knowledge to bear on design problems. To generate data, a number of expert and novice engineering students were asked to design a playground. Each subject’s behavior was tracked and categorized, but not for the purpose of describing cognitive processes. Instead the frequency of certain types of behavior between the two groups was compared, as was the content of the resulting designs and the strength of correlations between design quality and different types of behavior. One finding resulting from the study is that knowledge seeking activities correlate with design quality for both novices and experts.

A variation on verbal analysis that has been used to study knowledge representations involves analyzing drawings produced by subjects rather than verbal responses. By looking at resulting designs, for example, some researchers have attempted to infer both the kinds of knowledge and the structure of the knowledge used to solve design problems. Similar strategies have been used in studies of creative concept generation in cognitive psychology (e.g. Ward 95), as well as in developmental studies (e.g., Karmiloff-Smith 90). In design, Jansson and Smith (1991) use the strategy to study how the presentation of old solutions to a product design problem constrains the generation of new solutions. They found that attributes and structural relations possessed by the old solutions were more likely to be included in new solutions when the old solutions were presented beforehand, even when subjects were told that the old solutions are flawed and should not be copied. They concluded that designers tend to structure their solutions
on available exemplars rather than on a priori principles and seek novelty by varying attribute values rather than structural relations.

4. Process isolation

Content analysis attempts to overcome the limitations of think-aloud protocols by concentrating on the content of thought, particularly on the kinds of knowledge representations used in design. It does so, however, at the expense of understanding the processes that mediate representations. A second alternative to think-aloud protocols that aims at cognitive processes as well as at representations, involves the controlled study of processes isolated from drawn-out design episodes. This strategy assumes that a host of different activities undertaken by designers – not just conscious reasoning strategies – may contribute to the transformation of representations in design, leading, for example, to the generation of new problem representations and new ways of structuring knowledge about artifacts. These activities include those traditionally studied in cognitive science, such as analogy (e.g., Gick and Holyoak 83), mental modeling (e.g., Johnson-Laird 83), simulation (e.g., Barsalou 99), conceptual combination (e.g., Wisneiwski 97) and conceptual blending (e.g., Fauconnier and Turner 98).

Studies focusing on isolated processes are free to concentrate on fundamental cognitive theories apart from design provided that such theories can ultimately be combined into higher-level accounts of complex problem solving. More often, though, design researchers who focus on isolated processes tend to address the question of how existing theories in cognitive science can be extended to cover design behavior. Some design researchers, for example, have converted experimental tasks commonly used to study isolated cognitive processes into tasks that more closely resemble design or have picked up on experimental tasks that already resemble design (e.g., tasks used by Gestaltists, like the hat rack problem in Maier, 30). The result is often an experimental task that mimics design in some way but that reduces its complexity to allow for the controlled manipulation of independent variables.

4.1 Process isolation in design

Examples of process isolation studies in design include Akin and Akin (1996), Casakin and Goldschmidt (1999), Lawson (1979), and Verstijnen et al. (1998). In Akin and Akin (1996), processes associated with insight are examined using standard problems from the cognitive science literature (e.g., the nine-dot problem) as well as simple design tasks like the composition of an elevation given a small number of fixed constraints. In Casakin and Goldschmidt (1999), analogy is studied by giving designers simple design tasks lasting 10 to 15 minutes while showing them a variety of visual stimuli. In both studies, the tasks given to subjects include controlled miniaturizations of full-blown design problems. In other studies, more abstract tasks, which are taken to be representative of some part of design problem-solving, are used instead. In Verstijnen et al. (1998), for example, emergence in mental imagery and in drawings is studied using simple patterns and pre-functional shape compositions. Finally, in Lawson (1979), a geometrical puzzle is given to both scientists and designers to see how they compare in making inferences about problems and potential solutions.
The trick in studying isolated processes is coming up with experimental tasks that have well-defined dependent variables that are still relevant to design. The more a task is constructed to resemble design, however, the more murky dependent variables may become. In Casakin and Goldschmidt (1999), for example, the general quality of designs drawn by subjects, as assessed by a panel of experienced designers, was used as the dependent variable. The problem, in this case, is that design quality, although obviously of interest to design researchers, is not a unitary variable. In Casakin and Goldschmidt’s study, design quality may be confounded with variables other than those directly related to the analogical processes being studied. And although the finding that design quality is improved by the presence of visual analogs even when subjects are not told to use them is striking, it is limited in the extent to which it elucidates the underlying cognitive processes. Such studies might be expanded by incorporating content analysis, as was done in the study by Atman et al. (1999) discussed earlier, to extract finer-grained dependent variables that are relevant both to design and to the specific processes under investigation.

5. Situated studies

A final strategy for studying design behavior involves deferring theories of cognitive processes and representations altogether and instead focusing on design activities as they relate to their social, cultural and material contexts – in other words, treating the designer and the social, cultural and material contexts as “intact activity systems” (Greeno 98). In favor of such studies, a variety of researchers in different areas have argued that understanding situated behavior is essential for framing research on lower-level behavior and cognition, much like an understanding of behavior and cognition is essential for framing even lower-level neural theories (Greeno 98; Hutchins 95; Lave and Wenger 91; Suchman 87). These researchers, moreover, call for the development and use of theoretical concepts specific to situated behavior. They argue, in other words, that theories of situated behavior must be developed independently if they are to have a top-down influence on lower-level theories. Working cognitively from the bottom up may be seen as insufficient for two reasons. First, low-level processes might interact in unpredictable ways when integrated in more complex behavior, something that is overlooked in the study of isolated low-level processes. Second, low-level research is selective, in that it is framed by common-sense notions of situated behavior; hence, it may overlook processes that, in fact, play a significant role in reasoning. Low-level research framed by more considered investigations of situated behavior will, of course, also be selective, but presumably in a more insightful way.

Studies of situated behavior typically focus on one of two things: on the way meaning is produced in situations or on the way the social context and material environment regulate behavior. Situated studies focusing on meaning start by pointing out that meaning is not literally carried in external or internal symbols, but rather is constructed in particular situations. Generally, meaning is tied to a symbol’s referential status, and it is often through interaction with others and with the environment that one is able to establish (both for the agent and those he or she might be interacting with) the referential context, which includes, at various levels of abstraction, both what is being referred to and the perspective from which it is being referred. On the one hand, this
suggests that meaning may be constructed within and for specific situations, arising in a "process of co-construction and negotiation between participants and other systems" (Greeno 98, p. 14). At another level, it suggests that a plan generated in the head (however it happens be externalized) may have a limited significance in studies of behavior since the meaning of a plan ultimately takes shape in a situation (Suchman 87). To put it simply, people execute plans in action; plans do not execute themselves.

Understanding meaning, of course, is not an end in itself. Meaning ultimately underpins the reasoning practices and continued social and material interactions that are the aim of traditional cognitive studies. This, however, does not mean that situated studies end when the person being studied retreats to a private office to work alone. Problems faced by individuals get their meaning, both as specific problems and as instances of general types of problems, from socio-cultural situations that extend beyond the office, as do the concepts with which people reason. Presumably, if one wants to better understand the sorts of concepts an individual is using regardless of his or her immediate context, studying how they arise through interaction with intact activity systems will yield insights that pure cognitive studies fail to generate.

Situated studies that focus on the way the social context and material environment regulate behavior take a similar approach to studies that focus on the production of meaning. Here, the primary assumption is that social situations and material environments afford certain types of behavior. Just like a cocktail party, for example, affords certain types of behavior, so does a design office or school classroom. In part, individuals become effective practitioners by learning how to participate (Lave and Wenger 91) or, in different terms, how to take advantage of social affordances (Greeno 98). Individuals may also manipulate the environment in order to produce material affordances specific to a given task (e.g., Kirlik 98). Researchers in this area have focused on the structure of social and material affordances as well as on how individuals interact with and reproduce them. This type of research is seen as particularly crucial if the goal is to improve situations, particularly situations associated with education and professional practice. If, by contrast, one develops a general cognitive model and applies it to a specific situation, the social and material contexts may get in the way. The social and material contexts, moreover, cannot simply be removed to make room for ideal practices. Context, rather, is taken to be constitutive of practice. For this reason, one method suggested for studying situated behavior is to actively participate in social and material interventions to see exactly how they interact with existing affordances and how they foster new ones (Greeno 98).

**5.1 Situated studies in design**

Design studies like those discussed in the previous sections are in some respects naturally situated. For one thing, design problems generally take on meaning as they are being worked on – that is, in a situation. A designer may start off with a plan in mind, but plans are generally thought to be of limited value in solving design problems (Ullman et al. 88), especially when compared to well-defined problems. In this sense, design researchers naturally extend the boundaries of traditional cognitive studies when they study design. Additionally, design studies often aspire to some degree of realism.
Subjects in protocol studies, for example, are often given a range of realistic resources, sometimes including office environments, artifacts, experts and even teammates. Design problems are also often presented in a realistic fashion. If, for example, a problem used in a protocol study must be worked out in an unusually short amount of time, subjects may be told to imagine their task as one of preparing quick concept sketches for a rushed meeting with the client.

Despite the naturally situated nature of some design studies, several factors limit the extent to which experimental settings capture the situated nature of actual practices. First, experimental settings seldom include a persistent social network or a personally tuned physical environment, as one might expect to find in a design office or school. A typical office includes, among other things, coworkers with whom a designer has already developed social relations. Even if a design experiment involves multiple designers working together, in most cases subjects will enter the experimental situation as strangers. In addition, subjects may use concepts in experimental settings that depend on their everyday work environments for meaning. When removed from their everyday settings, these concepts may lose meaning.

Another problem with design studies conducted in experimental settings is that subjects may have difficulty constructing a meaningful image of the client (Lloyd et al. 96) or interactively negotiating meaning with them (although see Dorst, 97, for an example of how a subject’s view of the client may take shape in a think-aloud protocol). Finally, the social situation in studies of individual designers usually centers on the experimenter and hence offers few affordances for interaction. Although subjects in think-aloud protocols may interact with the experimenter, the experimenter is usually equipped to answer questions in a scripted manner, precluding any sort of unstructured communication. In situations outside the lab, however, casual interaction may actually play an important role in reasoning about tentative, speculative ideas (e.g., Roschelle 92; Teasley 94).

While experimental studies are generally lacking in their treatment of situated behavior, there are a handful of studies that do focus specifically on situated aspects of design. Some design studies, both in situ and in the lab, have looked at the role of dynamic interaction with the environment and with other participants. Goel (1995), for example, looks at the way interaction with sketches helps facilitate lateral transformations as subjects generate design alternatives. Similar studies dealing with the material environment include Harrison and Minneman (1996), which looks at how artifacts in the immediate environment are used improvisationally in team processes to illustrate and reason about alternatives, and Carter (1993), which looks at how an architect arranges drawings around his desk to facilitate the retrieval of information within a smooth flow of activities. Carter’s observations, as well as those of Minneman and Harrison, and Goel, are consistent with studies in cognitive science that highlight purposeful manipulations of the environment aimed at making reasoning easier (e.g., Kirsch 95; Hutchins 95). Studies focusing on interaction with other people, on the other hand, include Schon’s (1982) study of a student interacting back and forth with an expert designer in an actual studio setting. From his observations Schon argues that
Knowing and Learning to Design

such interaction helps bring new information into the picture and resolve the designer’s understanding of the problem.

A variety of studies in design have also addressed social interaction within design teams, some in actual work settings (e.g., Bucciarelli 94; Lloyd and Deasley 98), others in experimental settings (Cross and Cross 96; Berenton et al. 96). Such studies tend to focus on how the behaviors of different individuals contribute in different ways to a group problem-solving process. Some team members might, for example, be observed organizing and collecting votes on design alternatives, while others are observed balancing this sort of behavior by striving to keep design options open. These sorts of studies speculate on group problem solving independent of studies on individual problem solvers. In other words, they tend to treat groups as unique types of cognitive systems. Interaction among team members, for example, may be analyzed in terms of how it overcomes problems like fixation that individuals working outside a group often succumb to (cf. Hutchins 91).

Few studies in design deal specifically with how concepts are constructed, reproduced or conferred through local practices. Most seem to assume that although knowledge may be subjective, it is ultimately derived from experiences via general reasoning (e.g., inductive generalization or pattern recognition). Schon (1984), however, discusses the use of design concepts specific to local practices, in particular the role certain exemplars and metaphors play in the architectural program at MIT. He discusses, for example, the nuances and tensions associated with the use of references like “spaghetti bowl” and “Renaissance order.” Although he observes such references being used in the architectural studio, he stops short of investigating the broader culture in which they were developed. Thus, one is left with a limited understanding of their content and their place within the institution.

6. Summary

Four distinct strategies for studying design behavior have been discussed in this chapter. Although the strategies were not presented in any particular order, they naturally cluster into two contrasting pairs. Protocol analysis and content analysis, on the one hand, place contrasting emphasis on the transformational moves and representations used by designers. Although they appear more competitive than complementary, both encounter their share of difficulties when applied to design. In particular, while it tends to be difficult to track transformational moves in an ill-defined problem space, it is similarly difficult to uncover knowledge representations when a subject is free to draw on a potentially limitless set of related knowledge.

Process isolation and situated studies, on the other hand, take contrasting top-down and bottom-up approaches to studying design behavior. Although they also appear more competitive than complementary, it is not clear that one outperforms the other. Although process isolation may benefit from the sleuth work of situated studies, situated studies may also be informed by process isolation. Moreover, process isolation possesses a quality that has been fundamental to the success of science since its
inception yet is lacking in the other strategies: the ability to control and measure different phenomena independently.

Admittedly, not many of the studies discussed in the chapter fit neatly into any one of the four categories that were discussed. Most refer, in fact, to Ericsson and Simon (1984) for methodological support but end up using a mix of strategies in pursuit of a mix of goals. Recognizing the multifaceted nature of most research programs, the criticisms laid out in this chapter are meant to be illustrative rather than indicting – serving, that is, to probe the strengths and weaknesses of the strategies rather than the individual studies themselves. At this point, however, what is clearly needed is more methodological clarity. Studies may be applauded for developing hybrid research strategies but not for articulating them. At a minimum, methodological clarity will boost existing studies of design behavior both by encouraging the continued refinement of research methods and techniques, and by helping researchers merge findings from complementary investigations.

References


Knowing and Learning to Design


Stalking Homo Faber


Knowing and Learning to Design


Nisbett R, and Wilson T, (1977) Telling more than we can know: Verbal reports on mental processes, Psychological Review, 84 (3), 231-259


Schooler J W, and Engstler-Schooler T Y, (1990) Verbal overshadowing of visual memories: Some things are better left unsaid, Cognitive Psychology, 22 (1), 36-71


Stalking Homo Faber


