

The Role and Impact of Mental Simulation in Design

BO T. CHRISTENSEN^{1*} and CHRISTIAN D. SCHUNN²

¹*Department of Marketing, Copenhagen Business School, Frederiksberg, Denmark*

²*Learning Research and Development Center, University of Pittsburgh, USA*

SUMMARY

Although theories of mental simulations have used different formulations of the premises of ‘thought experiments’, they can be fitted under a minimalist hypothesis stating that mental simulations are run under situations of uncertainty to turn that uncertainty into approximate answers. Three basic assumptions of mental simulations were tested by using naturalistic data from engineering design. Results from the design protocols showed (1) initial representations in mental simulation had higher than base-rate uncertainty, (2) uncertainty in mental simulations were lowered after simulation runs, (3) resulting representations had more approximations than base-rate or initial representations. Further, the reference to external representational systems (sketches and prototypes) was examined. It was found that prototypes had fewer technical/functional simulations compared to sketches or unsupported cognition. Although prototypes were associated with more approximation than unsupported cognition, the different external representation categories did not differ in information uncertainty. The results support the minimalist hypothesis of mental simulations. Copyright © 2008 John Wiley & Sons, Ltd.

In the domains of design and architecture, mental simulation has been anecdotally described as being of utmost importance. For example, one famous anecdote (Tafel, 1979) describes how Frank Lloyd Wright in 1928 developed the concept for Fallingwater, one of the most famous residential buildings in America. The building had been commissioned by Edgar Kaufmann who would telephone Wright to inquire how the plans were progressing. For 15 months the response was simply that the plans were proceeding well, although no drawings were made. One day Kaufmann called and proclaimed that he was en route—a mere 140 minutes away. At that point, Wright started sketching out the plans for Fallingwater; first floor plan, second floor, section, elevation, details. The plans were drawn up very near to their final form, apparently developed fully in Wright’s head prior to producing external representations. Wright later described his design process in the following manner:

‘Conceive the building in the imagination, not on paper but the mind, thoroughly—before touching paper. . . Let it live there—gradually taking more definite form before committing it to the draughting board. When the thing lives for you—start to plan it with tools. Not before. To draw during conception or sketch, as we say, experimenting with practical adjustments to scale is well enough if the conception is clear enough to be

*Correspondence to: Bo T. Christensen, Department of Marketing, Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark. E-mail: bc.marktg@cbs.dk

firmly held. . . But if the original concept is lost as the drawing proceeds, throw all away and begin afresh' (Wright, 1928).

This anecdote suggests that design concept development can, and perhaps should, be carried out by using only mental simulation and imagery, and that the concept perhaps should not be committed to external representations (paper or prototype) prior to a fully developed concept.

The present study tested whether mental simulations are run in design situations associated with uncertainty in order to turn that uncertainty into approximate answers. Theoretical accounts of 'thought experiments' are related to the three research paradigms; the first two concerning 'mental models', and the last to 'judgement under uncertainty'. We describe each below.

A mental model is a representation of some domain or situation that supports understanding, reasoning and prediction (Gentner, 2002). Craik (1943) was the first to propose a theory of mental models as dynamic representations or simulations of the world, but since then the 'mental models' research paradigm has not evolved into a unified field of theories and hypotheses (e.g. Forbus & Gentner, 1997; Markman & Gentner, 2001; Nersessian, 1992, 2002). Rather, a number of different theories each rely on different assumptions of the nature of the cognitive system and focus their research on somewhat different aspects of mental models. Mental models represent different interpretations of the possible relationship among objects and their properties based upon given information.

One approach focuses on how people perform logical reasoning tasks with mental models (Johnson-Laird, 1983). Here mental models are viewed as temporary reasoning structures held in WM that embody a representation of salient spatial and temporal relations among, and the causal structures connecting, the event, processes and operations of complex systems. This approach has been utilized to study amodal token tasks, such as classical syllogisms and inductive arguments (e.g. Johnson-Laird & Byrne, 1991). One finding is that the fewer possibilities that have to be thought about for an inference, the easier and less error-prone is the inference (Johnson-Laird, 2001).

Another approach centres on causal mental models that are used in reasoning, and are based on long-term domain knowledge or theories (Gentner & Stevens, 1983). The focus lies on mechanical reasoning and reasoning about physical systems (Hegarty, 1992; Hegarty & Just, 1993; Schwartz & Black, 1996), navigation (Norman, 1983) and the development of astronomical knowledge (Vosniadou & Brewer, 1992). Causal mental models have some disadvantages, notably inaccuracy and imprecision (Gentner, 2002).

Causal mental models rely on qualitative relationships, such as signs and ordinal relationships, and relative positions, speed or mass (e.g. Forbus, 1984; Forbus & Gentner, 1997; Kuipers, 1994). When running mental models, people typically do not estimate exact values or carry out mathematical calculations in predicting system behaviour. Still, these qualitative reasoning strategies can be quite powerful and have the tremendous advantage of allowing reasoning with partial knowledge.

An important feature of causal mental models is that they frequently permit mental simulation: being able to dynamically 'run' a simulation internally to observe functioning and outcome of a system or device. 'Runability' is the ability to simulate system behaviour and predict outcomes even for situations where the subject has no previous experience. The potential advantages of using mental model runs in design include being able to reason about how physical systems will operate under changed circumstances/with altered features without having to resort to physically constructing such a system or device. This

ability is particularly useful in creative domains, such as science, art and design, where uncertainty is an inescapable part of the problem space.

A third approach to the study of mental simulation is usually attributed to Kahneman and Tversky (1982), where they describe the 'simulation heuristic' as a specialized adaptation of the availability heuristics taking place under situations of uncertainty. This approach is usually not described as related to the above theories of mental models, but nonetheless focuses on mental simulation, albeit not on the actual simulation run. Tversky and Kahneman used the 'simulation heuristic' to explain counterfactual thinking and its connection to perceptions of causality and regret by arguing that people determine the likelihood of an event based upon the ease of picturing it mentally. Social psychological explanations of counterfactual thinking have built upon this notion of a simulation heuristic, focussing primarily on the simulations of alternatives to autobiographical episodes and the associated affect (see Roese, 1997; Roese & Olson, 1995). Counterfactual thought is believed to serve preparative functions through useful causal inference (Roese, 1997).

Although the three approaches to mental simulation differ widely in scope, focus and basic assumptions about the underlying nature of the cognitive system that supports mental simulation, they do have common elements linking them together. As Nersessian (2002) has argued concerning theories of mental models, although the issue of the nature of the WM and LTM representations involved in reasoning with mental models need eventually to be resolved, they do not need to be settled before it is possible to make progress on an account of model-based reasoning in particular domains, such as science or design. If a 'minimalist' hypothesis to mental simulation is adopted, it is possible to study mental simulation across all these theories. Here (following Nersessian, 2002, p. 143), we minimally hypothesize that humans reason by constructing a mental model of the situations, events and processes in WM that in dynamic cases can be manipulated through simulation.

Whereas a mental model is runnable, the mental simulation is the actual 'run' (Trickett & Trafton, 2002, 2007; Trickett, 2004; Trickett, Trafton, Saner, & Schunn, 2005), meaning that the mental simulation will be explicitly present in working memory and may therefore be articulated by the participant (Ericsson & Simon, 1999). Thus, protocol studies of mental simulation may provide concurrent verbal accounts of mental model runs, as opposed to the usual inferred usage of mental models that has been the focus of much previous research. Although previous researchers have examined mental models using protocols (Chi, Feltovich, & Glaser, 1981; Clement, 1989, 1994, 2006; Dunbar, 1995, 1999; Griffith, Nersessian, & Goel, 2000), only Trickett (2004; Trickett & Trafton 2002; Trickett et al., 2005; Trickett & Trafton, 2007) have used the present methodology of locating mental model run in transcripts. In their case, it was scientists doing data analysis. They found that mental simulations (referred to as conceptual simulations) were reliably detectable in the transcripts.

Using the minimalist hypothesis and verbal protocol measures of mental simulations, it becomes possible to test some of the basic assumptions behind theories of mental simulation. The present study tested the following three basic assumptions about mental simulations:

- (1) Mental simulations are more frequently run under situations associated with uncertainty. This assumption was explicit in the 'simulation heuristic' approach of Kahneman and Tversky (1982). Forbus (1988, 1997) has also argued that one of the core advantages of reasoning with mental models is that they accommodate ambiguity and

uncertainty. Trickett and Trafton (2002, 2007) hypothesized that mental simulations are most likely to take place in situations of uncertainty, and they found evidence for this in the scientific discovery process, in that periods of hypothesis evaluation (i.e. periods of uncertainty) were associated with mental simulation. Mental simulations were used as frequently as or more frequently than any other strategy, and thus played a significant role in scientists' consideration and evaluation of hypotheses. The present study tests more directly whether subjectively experienced uncertainty (measured as subjective exclamations of epistemic uncertainty, such as 'I don't know' or 'I'm not sure') is associated with the running of mental simulations.

- (2) Mental simulations of possibilities inform reality through inferences, reducing uncertainty. One key component to mental simulation theories is the assumption that mental simulations help gain knowledge of the actual world through simulations of multiple possible (either past or future) alternatives (Forbus, 1984; Spellman & Mandel, 1999). Knowledge is gained mainly through inferences (Gentner, 2002), leading to the generation of possible alternatives or predictions for future encounters with similar situations. Functionally, mental simulations may thus reduce information uncertainty. For example, Nersessian (1992) argued that scientists conduct mental experiments that yield usable provisional data in processes mimicking empirical experiments. Schwartz and Black (1996) showed that how subjects used depictive models to generate parity rules when solving gear problems. Roese (1997) argued that causal inference is one of only two mechanisms accounting for nearly all documented counterfactual consequences. Counterfactual thinking is thus one source of causal information that may be used by an individual in expectancies and intentions. In these and many more cases, mental simulation is associated with making inferences from mental experiments, which in turn can be said to reduce some of the associated uncertainty by leading to predictions or generation of new alternatives. However, none of these studies have directly tested whether mental simulations lead to lower subjectively experienced levels of uncertainty (i.e. epistemic uncertainty). The present study examines whether subjective levels of epistemic uncertainty are reduced through mental simulations. It is possible, however, that mental simulations, being entirely mental increase subjective uncertainty levels—that is, the subject knows mental simulations are creating uncertain inferences that must be verified through other means, thus increasing subjective uncertainty levels about the results of the simulation.
- (3) The result of the mental simulation is approximate (i.e. non-exact). Mental simulation has been called a qualitative reasoning strategy since it does not involve quantifications of alternatives, but can operate with inexact or missing information, leading to partial, imprecise or approximate results (Forbus et al., 1997; Kuipers, 1994). So, even though mental simulation can be said to be biased and error-prone in some respects, it is still a powerful reasoning strategy as it allows for the quick and cheap generation of approximate or imprecise knowledge in situations where the production of exact quantitative simulations is inappropriate, unavailable or impossible. However, this 'approximate' nature of the results is usually just assumed. The present study directly tested whether the result of mental simulations can be said to be turning uncertain information into approximate (non-exact) knowledge. Although approximation and uncertainty may seem to be similar terms, here they are closer to opposites. Approximate answers operate within boundaries to some quantitative degree (as in 'the answer will be *around* X'). Even though approximate statements do not give an exact answer, they are not uncertain in the epistemic sense, that is, they are certain within boundaries.

For example, the informational uncertainty question, 'How old are you?' can result in an approximate but certain response, '30-something'. Alternatively to mental simulations producing approximations, it is possible that mental simulations primarily result in qualitative outcomes with no approximate quantifications.

Although these three assumptions are implicit in most theories of mental simulation, their overall function has not been specifically described as turning uncertainty into approximate answers. Further, these assumptions have not been tested directly in experiments or in naturalistic settings. We examined each of these three basic assumptions in a naturalistic (in vivo) setting, using protocol data from engineering designers.

EXTERNAL REPRESENTATIONS AND DESIGN

A second core goal of this study is to unpack external factors that evoke or influence the use of mental simulations. Mental simulations do not always occur in a vacuum. Sometimes the problem solver does everything mentally. But, more often, the mental simulations are done in the context of a physical object or drawings. Parts of the objects are mentally modified or transformed, and the external objects serve as both a help (providing a very concrete base situation) and a hindrance (interference between the perceived world and the imagined world).

It is a little explored avenue in cognitive science how creative cognitive phenomena interact with the presence or absence of external support systems. In the literature, such interaction studies have typically been limited to lab experiments contrasting sketching with no sketching in order to examine characteristics of imagery, such as the possibility of restructuring (e.g. Anderson & Helstrup, 1993; Finke, 1990; Verstijnen, van Leeuwen, Goldschmidt, Hamel, & Hennessey, 1998; Verstijnen, van Leeuwen, Hamel, & Hennessey, 2000). Even less studied than sketching, however, is cognition with the external support of physical objects, such as prototypes. Prototyping and sketching share creative properties with unsupported cognitive structures, in that they are '... formed without full anticipation of their resulting meaning and interpretation. In addition they are to be distinguished from the final externalized creative products, which, in contrast, are often fully interpreted and extensively refined' (Finke, Ward, & Smith, 1992, p. 22–23). In concept design, prototypes may change appearance rapidly or be reinterpreted in terms of functional inference. As such, a prototype is both an object frozen in time and form and a 'fluid' structure able to change meaning and appearance.

However, despite their similarities, prototypes provide a rather different kind of support to the cognitive system than do sketches. It has been suggested that graphical representations may offer 'computational off-loading' through providing direct perceptual recognition and 'read-off' of elements and relations (Scaife & Rogers, 1996). While sketches support visual perception well, they do not support other sense modalities (haptic, gustatory, olfactory and auditory) as well as prototypes do. This difference would suggest that prototypes would be superior to sketches in terms of support for perceptual read-off in some (but not all) sense modalities. The ability to do perceptual read-off from sketches and prototypes should lead to lower degrees of uncertainty and more approximations when compared to ideas.

The cognitive support for perceptual read-off in prototypes may also lead to less need for mental simulation, when compared to sketches or unsupported ideas, since perceptual read-off is a less cognitively expensive strategy of information uncertainty reduction than is

mental simulation. Previous research on strategies for conducting data analysis using visual displays in science has demonstrated that perceptual read-off is much more frequently employed as a strategy when compared to mental simulation (Trickett et al., 2002).

A competing, but also plausible, hypothesis to the one advanced here is that the availability of the perceptual read-off in prototypes and partly in sketches could lead to an increase in the number of mental simulations for two reasons: (1) inherent design flaws may easily be perceptually picked up, and rapidly changed in mental simulations, or (2) affordances of prototypes may facilitate mental simulation better than sketching or having no external representation.

Mental simulation in design may be of at least two different types. Either the simulation can concern mechanical or functional simulations of the product, or the simulation may concern simulating end-user behaviour or usability of the object in question. In design, prototypes are believed to support primarily mechanical or functional mental simulations of product features. Therefore the hypothesized reduction in mental simulations for prototypes may be limited to these mechanical or functional mental simulations.

To test these hypotheses on mental simulation, external representational systems and their relation to epistemic measures of information uncertainty, we collected naturalistic data from engineering designers doing design problem solving.

METHODS

Research setting

Data were collected by using Dunbar's (1995, 1997) and Dunbar and Blanchette (2001) *in vivo* methodology. The *in vivo* methodology allows the researcher to study expert thinking and reasoning 'online' in the real world. To this end, a major international company focusing on engineering design of medical plastics was selected for its persistent creativity. The company's R&D department had won multiple design awards for a number of different designs. For reasons of confidentiality, the company will not be named here, and the exact nature of the design project will not be disclosed. A particular design project spanning 2 years was chosen as the focus of the current study. The design project involved a total of 19 expert engineering designers organized in three subgroups focusing on different aspects of the design object, and involved designing a new and improved product. The three subgroups were organized as multidisciplinary teams, involving different functions.

Each subgroup in the project held product development meetings on a weekly basis. Because the designers were talking out loud, there was an external record of thinking and reasoning. The primary function of these subgroup product development meetings was creative development of design artefacts—actual creating and problem solving in collaboration. The activities included brainstorming, concept development, design problem solving, planning of data collection and the next steps of the design process, testing and evaluating mock-ups and prototypes, sketching activity, conducting experiments and discussions and knowledge exchange about end-users and production methods. The vast majority of the design activity at these meetings concerned design thinking and reasoning in the here-and-now.

The current paper examines the meetings of one particular subgroup with the task of developing completely novel features for the new product. It consisted of five core

members (1 female and 4 male) representing different functions (industrial designer, lab technician, project manager) and backgrounds (machine engineering, architecture, machinist). They all had extensive experience in medical plastics and design (ranging 10–35 years).

The subgroup product development meetings were videotaped by using a camera capturing design objects present on the table between the designers and object handling (holding prototypes or sketching activity), albeit not in detail. Further, gestures and general direction of gaze of the designers could be discerned from the video. During the meetings, the experimenter was present as an observer only. No special instruction to think-aloud was given—the participants were merely asked to continue with the meeting as they normally would. Each meeting lasted between 30 minutes and 2 hours. The meetings were recorded during the concept phase or first 5 months of the design project. The designers would work on several different design concepts at each meeting, although usually two or three would be the main focus of each session.

Protocol segmentation and reduction

The recordings were transcribed and segmented according to complete thought (see e.g. Christensen, 2005; Hughes & Parkes, 2003). Segmentation entails separating verbal statements into segments containing verb phrases that are indicative of mental operations. Further, segments would also change with changes in who was talking. A total of seven transcripts covering approximately 9 hours of video (7773 segments) were used in the present study. The transcripts were initially reduced by coding for off-task behaviour (e.g. jokes, banter between the designers, office gossip or events unrelated to design) and episodes dealing with summarizing past meetings or planning future meetings or data collection, removing 1602 segments from further analysis.

Coding of type of external representational system

Each segment was supplemented with information regarding which (if any) design object presently in the room was the focus of attention of the person speaking (typically a sketch or prototype located on the table between the designers). This information was coded by using the video recording of the design session. Focus of attention was operationalized as either actual handling or holding a particular object, pointing to a particular object or gazing toward a particular object.

For each segment, we coded whether the verbally referenced design object was unsupported by external representations ('idea'), supported by sketches or supported by 3D physical objects in the form of prototypes. These types of design objects in-the-making will be referred to as 'external representation systems'. These three types exhausted the types of design objects in-the-making found in the transcripts. In addition, a 'finalized existing product' ('other') category was used to capture references solely to existing products in the market. These 'other' segments were excluded from further analysis because they were not considered creative design objects in-the-making.

Coding of mental simulation

Mental simulations codes were adapted from Trickett's (2004; Trickett & Trafton, 2002, 2007) coding scheme of scientists running mental models during data analysis. A mental

model run is a mentally constructed model of a situation, phenomenon or object that can be grounded in memory or in a mental modification of the design objects currently present. The key feature in a mental simulation is that it involves a simulation ‘run’ that alters the representation to produce a change of state (Trickett, 2004). The simulation is not merely a question asked (e.g. changing features or functions of the design object); it also provides a kind of answer (e.g. will it work, how should it be produced). Mental simulations thus represent a specific sequence starting with creating an initial representation, running the representation (configural transformation where elements or functions are extended, added or deleted), followed lastly by a changed representation. These three elements (initial representation, run and changed representation) are not mutually exclusive and may occur in the same utterance/segment, although frequently they will cover several segments. Each segment was coded as ‘mental simulation’ or ‘no mental simulation’, along with the separate steps (see Table 1 for an example).

Coding for mental simulation types

Each complete simulation was coded for whether it pertained to (1) functional or technical aspects of the product (e.g. simulating changing the form, functions, features of the product), or (2) to simulating end-user behaviour in terms of habits, usability, purchasing behaviour and so on.

Coding of information uncertainty

Information uncertainty was coded using a purely syntactical approach. This approach was adapted from Trickett et al. (2005), who used hedge words to locate segments displaying uncertainty. These hedge words included for example words like ‘probably’, ‘sort of’, ‘guess’, ‘maybe’, ‘possibly’, ‘don’t know’, ‘[don’t] think’, ‘[not] certain’, ‘believe’ and so on. Segments containing these hedge words were located and coded as ‘uncertainty present’ if a scrutiny of the individual segment confirmed that the hedge word concerned uncertainty. All other on-task here-and-now segments were coded as ‘no uncertainty’ (see Table 2 for examples).

Coding of information approximation

Approximation was coded by first syntactically locating hedge words adapted from Trickett et al. (2005). These hedge words included words like ‘pretty much’, ‘virtually’,

Table 1. Example of a mental simulation

Step	Utterance
Initial representation	Could you add something so that you couldn’t close this thing because there would be something in the way when you try to fold this way...
Run	But if this thing goes this way, then it is in a position to allow the ear to enter... But then I just don’t know how it should be folded... ’cause if it is folded this way then it will come out here... then it should be folded unevenly some how... You should fold it oblique
Changed representation	It wouldn’t make any difference one way or the other. It would fold the same way, and come out on this side the same way

Table 2. Examples of information uncertainty using syntax

Utterance	Code
'Cause I'm <i>not sure</i> whether you would fold it around the back	Uncertainty
I <i>think</i> so too, but before we get too cocky, let's make a model. . .	Uncertainty
Well, I <i>guess</i> it's a combination of moist and heat isn't it? I <i>suppose</i> it has to be	Uncertainty
It has to push from the start	No uncertainty
Yes, but the problem is that you can't hit it later. . . 'cause its too small	No uncertainty
It... then we have... then we loose the possibility of folding it back	No uncertainty

'generally', 'frequently', 'usually', 'normally', 'basically', 'almost' and so on. Segments containing these hedge words were then coded as 'approximation present' if a scrutiny of the individual segment confirmed that the hedge word concerned approximation. All other segments were coded as 'no approximation'.

RESULTS

Inter-rater reliability

The first author coded all transcripts. Reliability coding was done on two full transcripts (approximately 1200 segments or 18% of the data) by a second independent coder unaware of the hypotheses being tested. The second coder received training both in doing protocol analysis in general, and in applying the present coding scheme by using a subset of the data. All disagreements between coders were resolved through discussion. Reliability kappas are reported in Table 3.

To validate the distinction between uncertainty and approximation, the adjacency relationships of approximation and uncertainty codes were examined. If the coding of uncertainty and approximation were valid, then one would expect that uncertainty segments would be more likely to follow uncertainty segments and approximation segments to follow approximation segments, with the idea that psychological approximation and uncertainty tend to linger. Figure 1 shows that this expected relationship is indeed found, with the base-rates of approximation ($M=4\%$) and uncertainty statements ($M=5\%$) shown in dashed lines. Uncertainty was followed by uncertainty more frequently than base-rate (one-sample $t(377) = 5.9, p < .001$) and more frequently than approximation (paired $t(377) = 6.4, p < .001$). Conversely, approximation was followed by approximation more frequently than base-rate (one-sample $t(329) = 2.8, p < .01$) and more frequently than uncertainty (paired $t(329) = 3.1, p < .002$). Neither

Table 3. Inter-rater reliability Kappas for each coding dimension

Code dimension	Kappa
Type of external representational system	.73
Mental simulation	.90
Mental simulation type	.74
Information uncertainty	.95
Information approximation	.96

All codes reached a satisfactory level ($>.60$).

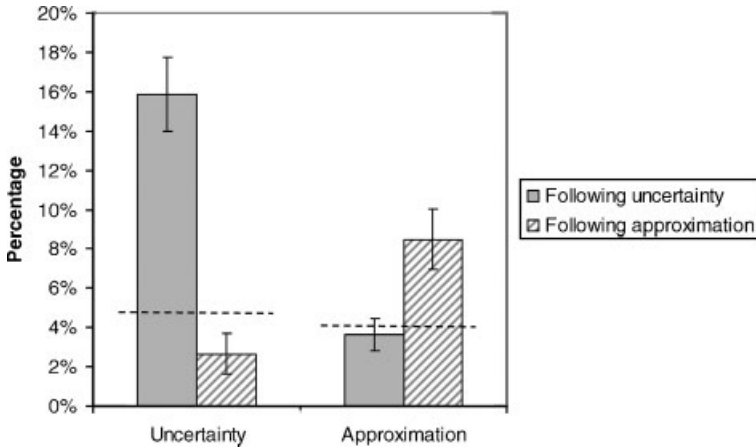


Figure 1. Percentage of segments with uncertainty and approximation following a segment with uncertainty or approximation with SE bars. Base-rates are shown with dotted lines

uncertainty following approximation nor approximation following uncertainty differed significantly from base-rate.

Frequency and form of mental simulations

A total of 327 complete simulations were found in the seven transcripts ($M=36$ simulations per hour of verbal data). The range of complete simulations was 20–87 per transcript. Simulation segments covered 33% of the on-task here-and-now segments in the present data set, showing that simulations were very commonly used by the designers during product development meetings (i.e. once every couple of minutes).

Of the 327 complete simulations, 68% contained distinguishable segments for all three sequential steps in the simulation: initial representation, simulation run and simulation result. The remaining simulations did not have segments dealing exclusively with one or two of the sequential steps. The number of segments in each complete simulation ranged from 1 to 22, with a mean of 3.6 segments per simulation. Thus, mental simulations varied in complexity. The two types of mental simulation both occurred frequently, with 218

Table 4. Examples of types of mental simulation

Type of simulation	Utterance
Functional/ technical simulations	‘But but but... you could imagine that they could make the filter that way so that you could use it... maybe you could use the filterdot also... it’s sitting right here [points to prototype]... but that would ruin the filterdot’
End-user simulation	‘Someone who knows... maybe... that they don’t use that much output... I mean when they’re not having output but wants to empty it... they would be able to carry it under their bathing suit... or something... folding it up or something... In that situation it would be of use’

counts of technical/functional simulations, and 109 end-user simulations (see Table 4 for examples).

Frequency of external representational systems

A total of 1842 segments referred to ideas unsupported by external representational systems, 235 segments referred to sketches and 1497 segments referred to prototypes.

The temporal relationship of mental simulation to uncertainty and approximation

To examine whether mental simulations were turning information uncertainty into approximation, we first compared the proportion of uncertainty segments in each sequential step within the mental simulations using paired *t*-tests. Results indicated that uncertainty in the initial representation ($M = 17\%$) differed significantly from uncertainty in the resulting representation ($M = 12\%$) ($t(292) = 2.1, p < .04$), with more uncertainty in the initial representation. The simulation run did not differ significantly from either the initial representation or the resulting representation (see Figure 2). The base-rate of uncertainty in segments not dealing with mental simulations was 8%. *T*-tests against the base-rate revealed that both the initial representation ($M = 17\%$) ($t(309) = 4.2, p < .001$) and the simulation run ($M = 13\%$) ($t(247) = 2.3, p < .03$) were more uncertain than base-rate, whereas the resulting representation ($M = 11\%$) uncertainty level was not above base-rate.

To examine whether mental simulations were leading to more approximations, we compared the proportion of segments with approximations in each mental simulation step using paired *t*-tests. Results indicated that approximation in the initial representation ($M = 3\%$) differed significantly from approximation in the resulting representation ($M = 7\%$) ($t(292) = 2.0, p < .05$), with more approximation in the resulting representation. The simulation run did not differ significantly from either the initial representation or the resulting representation (see Figure 2). The base-rate of approximation in segments not dealing with mental simulations was 3%. *T*-tests against the base-rate revealed that the resulting representation ($M = 7\%$) ($t(307) = 2.5, p < .02$) had more approximation than the base-rate, whereas the initial representation ($M = 4\%$) and the simulation run ($M = 6\%$) were not more approximate than base-rate. These results lend support to the hypothesis that mental simulations reduce information uncertainty by turning it into approximate answers in design.

The relationship between external representational systems and mental simulation

If mental simulation is used as a strategy to reduce information uncertainty, then different kinds of external representational systems may affect the need for mental simulations due to differing levels of associated uncertainty with ideas, sketches and prototypes. A 3 (external representational systems) \times 2 (presence/absence of mental simulation) chi-square revealed significant differences in the number of mental simulation segments in the three external representational systems categories ($\chi^2(2) = 201.0, p < .001$). Follow-up 2×2 chi-squares on each individual pair of external representational systems indicated that ideas had a significantly higher proportion of mental simulations than

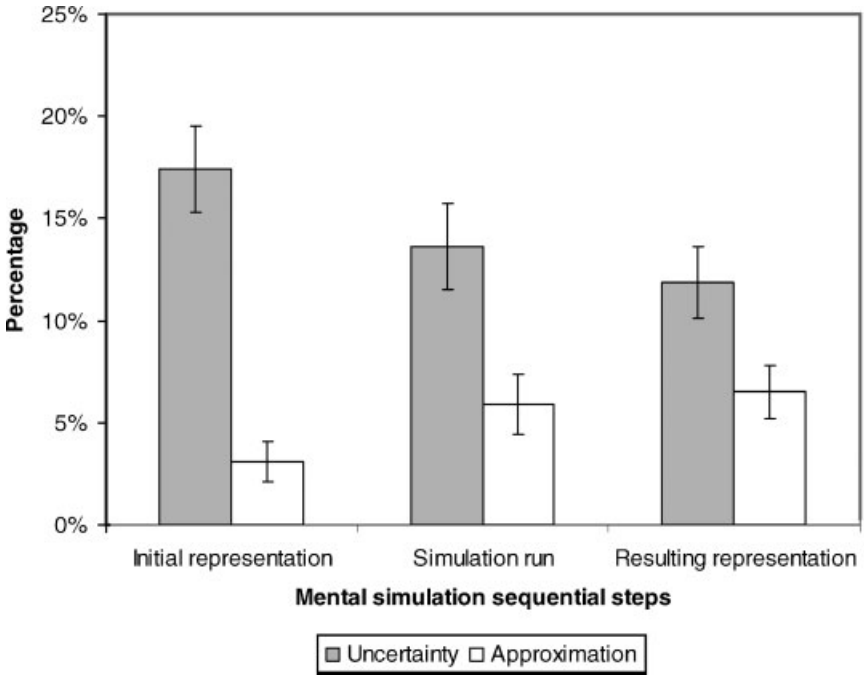


Figure 2. Percentage of segments with uncertainty and approximation by mental simulation sequential step with SE bars

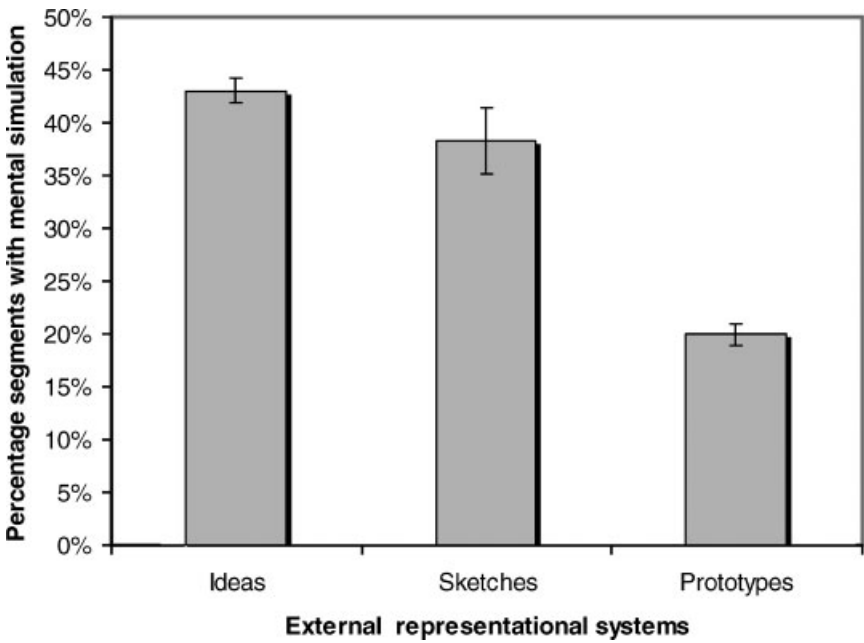


Figure 3. Percentage of segments with mental simulation by external representational systems with SE bars

prototypes ($\chi^2(1) = 199.0, p < .001$), sketches had a higher proportion of mental simulation than prototypes ($\chi^2(1) = 39.2, p < .001$), while ideas did not differ significantly from sketches (see Figure 3).

The result that prototypes are associated with less mental simulation could be caused by prototypes having less associated technical or functional uncertainty. A 3 (external representational systems) \times 2 (type of simulation) chi-square indicated that there were significant differences in the types of mental simulations between the external representational systems categories ($\chi^2(2) = 14.3, p < .001$). Subsequent 2×2 chi-squares conducted on each individual pair of external representational systems found that ideas had significantly more technical/functional simulations relative to end-user simulations when compared to prototypes ($\chi^2(1) = 11.9, p < .001$), as had sketches when compared to prototypes ($\chi^2(1) = 6.2, p < .013$), whereas ideas did not differ significantly from sketches (see Figure 4). This result suggests that the reason for the lower number of mental simulation segments for prototypes compared to ideas or sketches (noted above) may be the higher proportion of technical/functional mental simulations for ideas and sketches compared to prototypes.

The relationship between external representational systems and uncertainty/approximation

Constructing external representations may be another way of reducing uncertainty and increasing approximation. A chi-square conducted for type of external representational systems \times approximation present/absent indicated differences between the three types of external representational systems ($\chi^2(2) = 8.0, p < .02$). Subsequent 2×2 chi-squares conducted by each pair of external representational systems indicated that ideas differed from prototypes ($\chi^2(1) = 8.0, p < .006$), but sketches did not occur often enough to have their approximation levels reliably separated from ideas or prototypes (see Figure 5).

To examine whether different external representational systems were associated with differing levels of uncertainty, a chi-square was conducted for type of external

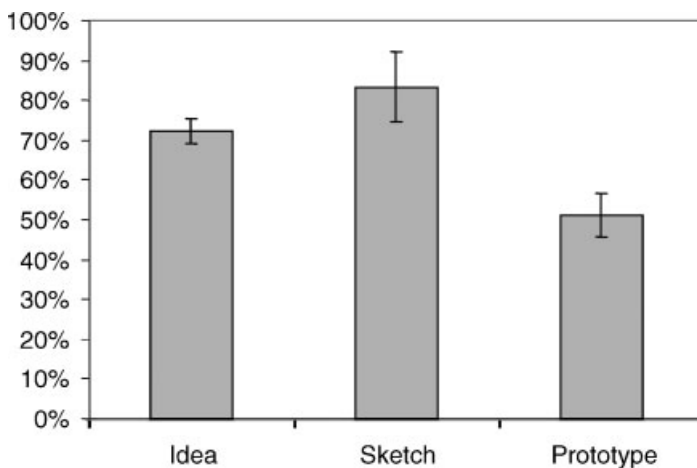


Figure 4. Percentage of technical/functional simulations of total simulations by external representational systems with SE bars

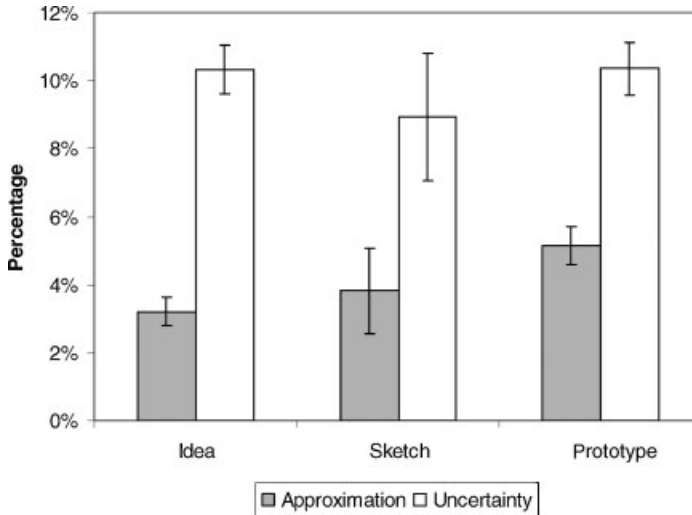


Figure 5. Percentage of segments with approximation and uncertainty by external representational systems with SE bars

representational systems × uncertainty present/absent. Surprisingly, uncertainty levels did not differ between the three types of external representational systems (see Figure 5).

Temporal development or order confound tests

In comparing ideas, sketches and prototypes in a naturalistic setting, it is important to consider whether there is an inherent temporal or order confound—can differences in performance with ideas vs. sketches vs. prototypes be attributed to *when* they are used in the design process. For example, unsupported ideas may be more prevalent in the early stages of problem solving, whereas prototypes may be more prevalent at the later stages, with sketches somewhere in between. This ‘natural’ time course of design artefacts, if true, might lead to confounds in the present results and spurious associations among the variables. To examine the issue empirically, all main variables were compared by temporal halves (ordered by data collection date). Neither mental simulation, nor uncertainty, nor approximation differed significantly between the first and second half. There was a small but statistically significant difference between type of external representational systems coded between the first and second halves. However, the difference appeared to be caused by slightly more prototyping occurring in the early category, and slightly more sketching in

Table 5. Mean occurrences of the main variables by temporal halves of the transcripts

	1st half	2nd half
Percentage of mental simulation	34	32
Percentage of uncertainty	10	11
Percentage of approximation	4	4
Percentage of idea	52	52
Percentage of sketch	3	11
Percentage of prototype	46	38

Numbers in bold differ significantly across temporal halves.

the latter category (see Table 5). As these differences are small and counter to what should be expected in a naïve understanding of the design process, the differences were not taken as evidence that temporal confounds explained the observed main phenomena.

GENERAL DISCUSSION

The current paper adds to a growing literature on the link between uncertainty and strategies utilized in design. For example, Ball et al. (Ball, Evans, Dennis, & Ormerod, 1997; Ball, Maskill, & Ormerod, 1998) have presented evidence that uncertainties about current solution ideas cause major strategic switches in the expert designer's ongoing problem solving (e.g. changes from optimal breadth-first to novice-like depth first or opportunistic design). Scrivener, Ball, and Tseng (2000) argued that uncertainty is the primary factor triggering change in drawing structure and a key driver of invention in design. They showed that uncertainty arises when internal representations are inadequate for the task at hand. At the same time, various researchers have argued that sketches are used (rather than prototypes or detailed technical drawings) because they are purposefully ambiguous, allowing for creative reinterpretation (e.g. Suwa, Tversky, Gero, & Purcell, 2001). Indeed our own prior work has showed that prototypes, but not sketches, appear to inhibit the generation of far-field analogies in the design team (Christensen & Schunn, 2007).

The present naturalistic data of engineering designers provide support to the basic assumption that mental simulations are run under situations of uncertainty, in order to turn that uncertainty into approximate answers. When mental simulations were split into initial representations, simulations runs and resulting representations, both the uncertainty in the initial and run parts of simulations were above the base-rate, attesting to the link between high levels of uncertainty and the running of mental simulations. Further, there was a significant reduction in uncertainty from the initial representation to the resulting representation in the mental simulations. The resulting representation did not differ from the base-rate. These results are consistent with an uncertainty reducing effect of running mental simulations. Mental simulations also appear to be producing approximate or non-exact results. Measures of approximation indicate that whereas the initial and run levels of approximation do not differ significantly from the base-rate, the resulting representation had significantly more approximations than both the base-rate and the initial representation rate.

In naturalistic data there is necessarily a great deal of noise and the lack of control of all variables usually makes it harder to find reliable effects. Moreover, syntactical approaches to coding schemes usually suffer from under-representing the number of instances compared to qualitative coding approaches. A qualitative scrutiny of individual segments would probably have increased the uncertainty and approximation levels overall, but perhaps at the expense of deflating inter-rater reliability. We found significant results using syntactic coding in a naturalistic setting, thus attesting to the robustness of the basic assumptions of mental simulation.

Although the minimalist hypothesis of mental simulations is supported by the present data, much theoretical and empirical work still needs to be done to contrast and synthesize the three different approaches to mental simulations mentioned in the introduction (concerning logical reasoning, causal mental models and simulation as a heuristic). How the WM structures are related to LTM and whether the LTM structure are also model-like

remains to be seen. The present study has illustrated that if mental models are seen as problem solving related reasoning processes in WM about situations, events and processes that in dynamic cases can be manipulated through simulation, then mental simulations appear to be linked closely to situations of uncertainty, turning uncertainty into approximate answers.

Another finding in the present study concerned the use of external representational systems (sketches and prototypes), and their relation to levels of uncertainty, approximation and mental models. The use of and reference to external representational systems in the form of prototypes appear to reduce the number of mental simulations in the present study, although this effect was limited to mechanical or functional simulations. Both cognition unsupported by external representational systems and cognition supported by sketches had more mental simulations than did cognition with the support of prototypes. This relationship supports the hypothesis that 3D external representational systems provide the option of using alternative strategies for reducing uncertainty, thereby limiting the need for running mental simulations.

Another possibility is that the construction of 3D external representational systems in and of themselves reduces uncertainty about the product. Further analysis showed that external representational systems were not related to lower levels of uncertainty when compared to unsupported cognition, although they were associated with more approximations. So, some support was found that external representational systems in and of themselves provided more approximate answers to uncertain questions in design. But we failed to find reduced levels of uncertainty for sketches or prototypes. There was, however, some anecdotal evidence in the present transcripts that the construction of prototypes were sometimes used to reduce technical or functional uncertainty. One good example was when the designers were simulating changing a design part radically. One design suggestion was so complicated technically, that even though the designers made several mental simulations over the course of 10 minutes, they were unable to figure out if the idea would work. Their mental model simulations would simply not generalize. They ended up disagreeing (and changing their minds several times) about whether or not it would work, and decided to make a prototype to figure it out. In this case, the designers did appear to decide to construct prototypes in order to reduce technical or functional uncertainty. However, our present coding scheme was unable to pick up any indication that prototypes overall are associated with less uncertainty compared to ideas. One possible explanation for this result is that the construction of prototypes itself reduces some technical or functional uncertainty, but may introduce new questions about other aspects of functionality. The above anecdotal evidence also illustrates that sometimes the mental simulation of ideas may not be enough to reduce uncertainty, and that producing sketches, mock-ups and prototypes is another (and perhaps secondary) strategy in this regard. Further research is needed to examine these questions, and to extend the generalizability of present results to other domains than engineering design.

ACKNOWLEDGEMENTS

This research was supported by a Carlsbergfondet grant to the first author and NSF grant BCS-0638451 to the second author.

REFERENCES

- Anderson, R. E., & Helstrup, T. (1993). Multiple perspectives on discovery and creativity in mind and on paper. In B. Roskos-Ewoldson, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and Discovery: A cognitive perspective* (pp. 223–252). Amsterdam: Elsevier Science Publishers B.V.
- Ball, L. J., Evans, J. S., Dennis, I., & Ormerod, T. C. (1997). Problem-solving strategies and expertise in engineering design. *Thinking and Reasoning*, 3, 247–270.
- Ball, L. J., Maskill, L., & Ormerod, T. C. (1998). Satisficing in engineering design: Causes, consequences and implications for design support. *Automation in Construction*, 7, 213–227.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.
- Christensen, B. T. (2005). A methodology for studying design cognition in the real-world. *Proceedings from the First Nordic Design Research Conference*.
- Christensen, B. T., & Schunn, C. D. (2007). The relationship between analogical distance to analogical function and pre-inventive structure: The case of engineering design. *Memory and Cognition*, 35, 29–38.
- Clement, J. J. (1989). Learning via model construction and criticism. In G. Glover, R. Ronning, & C. Reynolds (Eds.), *Handbook of creativity; assessment, theory and research* (pp. 341–381). New York, NY: Plenum Press.
- Clement, J. J. (1994). Use of physical intuition and imagistic simulation in expert problem solving. In D. Tirosh (Ed.), *Implicit and explicit knowledge* (pp. 204–244). Hillsdale, NJ: Ablex Publishing Corp.
- Clement, J. J. (2006). Thought experiments and imagery in expert protocols. In L. Magnani (Ed.), *Model based reasoning in science and engineering* (pp. 1–16). London: College Publications.
- Craik, K. (1943). *The nature of explanation*. Cambridge, UK: Cambridge University Press.
- Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365–395). Cambridge, MA, US: The MIT Press.
- Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 461–493). Washington, DC, US: American Psychological Association.
- Dunbar, K. (1999). How scientists build models: In vivo science as a window on the scientific mind. In L. Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 89–98). NY, US: Plenum Press.
- Dunbar, K., & Blanchette, I. (2001). The invivo/invitro approach to cognition: The case of analogy. *Trends in Cognitive Sciences*, 5, 334–339.
- Ericsson, K. A., & Simon, H. A. (1999). *Protocol analysis: Verbal reports as data* (3rd ed.). Cambridge, MA, US: The MIT Press.
- Finke, R. A. (1990). *Creative imagery: Discoveries and inventions in visualization*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Publ.
- Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.
- Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence*, 24, 85–168.
- Forbus, K. D. (1988). Introducing actions into qualitative simulations. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*.
- Forbus, K. D. (1997). Qualitative reasoning. In A. B. Tucker (Ed.), *CRC handbook of computer science and engineering*. Boca Raton, Florida, US: CRC Press.
- Forbus, K. D., & Gentner, D. (1997). Qualitative mental models: Simulations or memories? *Proceedings of the Eleventh International Workshop on Qualitative Reasoning*.
- Gentner, D. (2002). Psychology of mental models. In N. J. Smelser & P. B. Bates (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 9683–9687). Amsterdam: Elsevier.
- Gentner, D., & Stevens, A. (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum.
- Griffith, T. W., Nersessian, N. J., & Goel, A. K. (2000). Function-Follows-Form transformations in scientific problem solving. *Proceedings of the 22nd Cognitive Science Society Conference*.
- Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1084–1102.

- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717–742.
- Hughes, J., & Parkes, S. (2003). Trends in the use of verbal protocol analysis in software engineering research. *Behaviour and Information Technology*, 22, 127–140.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.
- Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences*, 5, 434–442.
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Peduction*. Hillsdale, NJ: Lawrence Erlbaum.
- Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgement under uncertainty. Heuristics and biases* (pp. 201–210). Cambridge, UK: Cambridge University Press.
- Kuipers, B. (1994). *Qualitative reasoning*. Cambridge, MA: The MIT Press.
- Markman, A. B., & Gentner, D. (2001). Thinking. *Annual Review of Psychology*, 52, 223–247.
- Nersessian, N. J. (1992). In the theoretician's laboratory: Thought experimenting as mental modelling. In D. Hull, M. Forbes, & K. Okruhlik (Eds.), *Proceedings of the 1992 biennial meeting of the Philosophy of Science Association*. East Lansing, Michigan: Philosophy of Science Association.
- Nersessian, N. J. (2002). The cognitive basis of model-based reasoning in science. In P. Carruthers & S. Stich (Eds.), *Cognitive basis of science* (pp. 133–153). New York, NY, US: Cambridge University Press.
- Norman, D. (1983). Some observations on mental models. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 7–14). Hillsdale, NJ: Lawrence Erlbaum.
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, 133–148.
- Roese, N. J., & Olson, J. M. (1995). *What might have been. The social psychology of counterfactual thinking*. Mahwah, NJ: Lawrence Erlbaum.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, 45, 213.
- Schwartz, D. L., & Black, J. B. (1996). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology*, 30, 154–219.
- Scrivener, S. A. R., Ball, L. J., & Tseng, W. (2000). Uncertainty and sketching behaviour. *Design Studies*, 21, 465–481.
- Spellman, B., & Mandel, D. R. (1999). When possibility informs reality: Counterfactual thinking as a cue to causality. *Current Directions in Psychological Science*, 8, 120–123.
- Suwa, M., Tversky, B., Gero, J. S., & Purcell, A. T. (2001). Seeing into sketches: Regrouping parts encourages new interpretations. In J. S. Gero, B. Tversky, & A. T. Purcell (Eds.), *Visual and spatial reasoning in Design II. Key centre of design computing and cognition* (pp. 207–219). Sydney: University of Sydney.
- Tafel, E. (1979). *Years with Frank Lloyd Wright: Apprentice to genius*. New York: Dover Publications, Inc.
- Trickett, S. B. (2004). *Movies-in-the-mind: The instantiation and use of conceptual simulations in scientific reasoning*. Unpublished Doctoral dissertation from George Mason University.
- Trickett, S. B., & Trafton, J. G. (2002). The instantiation and use of conceptual simulations in evaluating hypotheses: Movies-in-the-mind in scientific reasoning. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 878–883). Mahwah, NJ: Erlbaum.
- Trickett, S. B., & Trafton, J. G. (2007). 'What if': The use of conceptual simulations in scientific reasoning. *Cognitive Science*, 31, 843–875.
- Trickett, S. B., Trafton, J. G., Saner, L., & Schunn, C. D. (2005). 'I don't know what's going on there': The use of spatial transformations to deal with and resolve uncertainty in complex visualizations. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 65–85). Mahwah, NJ: Erlbaum.
- Verstijnen, I. M., van Leeuwen, C., Goldschmidt, G., Hamel, R., & Hennessey, J. M. (1998). Creative discovery in imagery and perception: Combining is relatively easy, restructuring takes a sketch. *Acta Psychologica*, 99, 177–200.
- Verstijnen, I. M., van Leeuwen, C., Hamel, R., & Hennessey, J. M. (2000). What imagery can't do and why sketching might help. *Empirical Studies of the Arts*, 18, 167–182.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the Earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535–558.
- Wright, F. L. (1928). Concept and plan. *The Architectural Record*, January–February.