Artificial Intelligence for Engineering Design, Analysis and Manufacturing

cambridge.org/aie

Speech analysis for conceptual CAD modeling using multi-modal interfaces: An investigation into Architects' and Engineers' speech preferences

Sumbul Khan[1] and Bige Tunçer[2]

[1]SUTD-MIT International Design Centre, Singapore University of Technology and Design, 8, Somapah Road, 487372, Singapore and [2]Architecture and Sustainable Design, Singapore University of Technology and Design, 8, Somapah Road, 487372, Singapore

Research Article

Cite this article: Khan S, Tunçer B (2019). Speech analysis for conceptual CAD modeling using multi-modal interfaces: An investigation into Architects' and Engineers' speech preferences. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **33**, 275–288. https://doi.org/10.1017/S0890060419000015

Received: 17 August 2017
Revised: 15 January 2019
Accepted: 17 January 2019
First published online: 14 March 2019

Key words:
Computational linguistics; computer-aided design; conceptual design; human–computer interaction; natural interfaces; speech analysis

Author for correspondence:
Sumbul Khan, E-mail: sumbul_khan@sutd.edu.sg

© Cambridge University Press 2019

Abstract

Speech- and gesture-based interfaces for computer-aided design (CAD) modeling must employ vocabulary suitable for target professional groups. We conducted an experiment with 40 participants from architecture and engineering backgrounds to elicit their speech preferences for four CAD manipulation tasks: Scale, Rotate, Copy, and Move. We compiled speech command terms used by participants and analyzed verbalizations based on three analytic themes: the exactness of descriptions, the granularity of descriptions, and the use of CAD legacy terms. We found that participants from both groups used precise and vague expressions in their verbalizations and used a median of three parameters in their verbalizations. Architects used CAD legacy terms more than Engineers in the tasks Scale and Rotate. Based on these findings, we give recommendations for the design of speech- and gesture-based interface for conceptual CAD modeling.

Introduction

In the conceptual design stage, designers conduct massing studies, which involve creating and manipulating three-dimensional (3D) forms (Akin and Moustapha, 2004). They study relationships between forms and their context. This is a creative design stage, wherein the design problem is ill-defined, and designers iterate possible solutions. Multi-modal interfaces using gestures and speech for computer-aided design (CAD) modeling are considered suitable for the conceptual design stage, as they offer an improved user experience and better control than conventional mouse and keyboard input (Oviatt, 1999). In this paper, we investigate the speech component for multimodal interfaces that employ speech and gestures in parallel, for conceptual CAD modeling.

Existing studies in speech-based input for conceptual CAD modeling are limited, and most employ vocabulary sets that are arbitrary or author-defined (Nanjundaswamy *et al.*, 2013), overlooking the specific needs of users in terms of the vocabulary used. Critics have argued that users should not have to learn an artificial language that is device or application dependent. They reason that users think and express in ways that cannot always be predicted (Malizia and Bellucci, 2012). A multi-modal interface is successful only if it is natural for its users (Quek *et al.*, 2002). The vocabulary employed in such interfaces must be natural for its users (Cassell, 1998). But what does a "natural" interface mean? Citing the example of gestural interfaces, Malizia and Bellucci (2012) define a natural interface as one with which people can interact with technology using the same gestures they employ with objects in everyday life, as conditioned by evolution and education.

Based on this view of natural interaction, we position that a multi-modal interface utilizing speech must be tailored to the vocabulary that specific user groups employ with the objects in their everyday professional life. It is widely acknowledged that the way people verbalize concepts about shapes and forms depends on their education, experience, socio-cultural, and linguistic background (Wiegers *et al.*, 2011). Previous studies have established that architects, owing to differences in education and training, have linguistic differences from other professional groups (Gifford *et al.*, 2002).

Hence, the goal of this study is to investigate the speech preferences of two groups of design professionals, architects and engineers, for conceptual CAD modeling using speech- and gesture-based interfaces. Architects and Engineers are the two primary user groups of CAD. We investigate whether both groups prefer to use precise expressions or vague expressions. Do engineers, owing to the emphasis of their study of math and sciences, employ more technical terms than architects, whose education is often more grounded in culture and aesthetics? Furthermore, present day designers are well versed with the WIMP (windows, icons, menus,

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

pointer) interface of CAD software, which has been in widespread use for the past few decades. Is the language employed by these groups influenced by their experience with the existing CAD software? Such investigations are necessary for the development of multimodal interfaces for conceptual CAD modeling that employ speech and are informed by user behavior.

To address these issues, we present an analysis from an experiment with 40[1] design professionals from architecture and engineering product development (EPD) backgrounds. The experiment was conducted to elicit speech and gesture preferences for conceptual CAD modeling from the stated professional groups. In related studies, we presented user preferences of gestures (Khan and Tuncer, n.d., 2017; Tunçer and Khan, 2018) and the implementation of a prototype (Khan *et al*., 2017).

In this paper, we present user preferences of speech terms for four CAD manipulation tasks, based on the professional backgrounds of the participants. First, we present relevant literature on the analysis of language employed by designers and the use of speech in virtual environments. We present a brief review of research on comparisons between architects and engineers ("Background"). Then, we elaborate the three analytic themes employed to analyze verbalizations ("Analytic themes for investigating verbalization"). In the following section, we describe our experimental design, coding scheme, and measures for the analysis of verbalizations ("Method"). In "Results", we present our results of the compilation of command terms, CAD legacy terms used by participants and analytical scores. In "Discussion", we discuss our findings on the exactness of descriptions, CAD legacy terms and the granularity of descriptions. Based on our findings and evidence from literature, we develop recommendations for the design of a multi-modal interface for conceptual CAD modeling. Finally, in "Conclusion", we summarize our chief findings and recommendations. We state the limitations of our research and future work.

This research analyzes aspects of natural language for the design of a speech- and gesture-based interface for conceptual CAD modeling. Due to its focus on human–computer interaction using natural language, the issues presented in this paper are pertinent to the domains of computational linguistics and, on scaling up, to natural language processing.

Background

Language is the method of human communication, either spoken or written, consisting of the use of words in a structured and conventional way. In this paper, we use the term "speech" to refer to spoken language, and "verbalization" to refer to speech strings, as used in previous studies (Wiegers *et al*., 2011). Our very thought and ideation process is dependent on language. Segers and Leclercq (2007) state that words are fundamental for communication and reasoning in the design process. In fact, studies stress that language is crucial for the very process of thought itself (Jonson, 2005).

Speech is an integral part of human communication. Speech is based on conventions and conveys meaning discretely, relying on codified words and grammatical devices (McNeill, 1992). Speech, along with gestures, is seen as a natural way to interact with computers (Mignonneau and Sommerer, 2005). Studies have investigated the role of language in designing activity (Lawson and Loke, 1997) and communication (Oak, 2011). Research has established the utility of speech-based interfaces for specifying spatial relationships (Bolt, 1980; Clay and Wilhelms, 1996). Furthermore, linguistic information enables user-friendly human–computer interaction that can handle interpretations and manage complexity in content (Segers and Leclercq, 2007).

Research has theorized that speech can be used effectively for descriptive tasks in the articulation of 3D form (Athavankar, 1999). Furthermore, restricted speech strings convey sufficient information about the form of the object (Varshney, 1998) and a few words are sufficient to capture the essence of products' semantic content (Lenau and Boelskifte, 2005). When free to interact multi-modally, users selectively eliminate linguistic complexities and employ briefer, syntactically simpler language (Oviatt, 1999).

Seminal studies in human–computer interaction have investigated how people describe objects to tailor information systems to specific audiences (Furnas *et al*., 1982). More recently, studies have investigated the words designers use to exteriorize shapes (Wiegers *et al*., 2011) and the issues related to the definitions of words in design discourse (Poggenpohl *et al*., 2004). Studies have analyzed designers' speech in experiments to recognize patterns, qualities, and quantities of speech (Purcell, 1996). Previous studies have also categorized designers' verbalizations in relationship to gesturing (Logan and Radcliffe, 2004).

Research into the use of multi-modal interaction for graphics system dates back to the 1970s (Brown *et al*., 1979). Multimodal interaction using gestures and natural language has been investigated for information retrieval and to provide system generated output (Neal *et al*., 1989). Studies on using speech as input for drawing and manipulating spatial objects include "Talk and Draw" (Salisbury *et al*., 1990) and Weimer and Ganapathy's (1989) speech and glove-based gesture input. Bolt (1980) investigated the use of speech with pointing gestures for the selection and displacement of two-dimensional virtual objects in the "Put-that-there" system. Investigating the use of imprecise speech input augmented with gestures, Bolt's study attempted to encapsulate natural human communication. The system proposed the use of commands such as "Create a blue square there" allowing users to employ vague language and use gestures for disambiguation. Speech has been used alongside gesture pen strokes as demonstrated in Herold and Stahovich's study (2011) in AIEDAM's special issue on the Role of Gesture in Designing. Recent studies using multi-modal input for CAD modeling include the studies of Menegotto (2015) who integrated speech with AutoCAD and Nanjundaswamy *et al.* (2013) who employed speech, gestures, and brain–computer interaction for invoking different CAD functionalities. Research in human–computer interaction has employed elicitation techniques to elicit gestures and speech interactions from users for diverse applications such as surface computing (Wobbrock *et al*., 2009) and web browsing (Morris, 2012).

Architecture and engineering professions both involve the design of artifacts such as those of buildings, products or automobiles. Yet architects and engineers have significant differences in education, training, and experience. Architecture education usually involves exposure to art and esthetics and is seen to be a creative field. On the other hand, engineering education involves a deep study of math and sciences and employs a more technical approach. Research has determined that architects perceive design artifacts and urban environment differently from other groups (Akalin *et al*., 2009; Ghomeshi and Jusan, 2013; Llinares and

---

[1] The experiment was originally conducted with 41 participants. However, since this part of the study compares preferences of architects vs. engineers, the data set of a randomly selected engineer was discarded to eliminate bias.

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

Iñarra, 2014). The differences in responses have been attributed to the different mental models or criteria employed by architects in their evaluations (Groat, 1982; Devlin, 1990) and their specialized training and exposure to studies of art (Berlyne, 1971; Llinares Millán *et al.*, 2018). Gifford *et al.* pointed out that architects have linguistic differences from other groups and base their evaluation on different sets of design features (Gifford *et al.*, 2002).

Current trends in both architecture and engineering education include collaborative approaches, the inclusion of technology and interdisciplinarity. Although boundaries within the field of engineering itself are blurring (Jørgensen, 2007), and a number of engineering departments have attempted to overcome the traditional division between civil engineering and architecture (Crawley *et al.*, 2014), recent studies still view architects and engineers as distinct groups (Najari *et al.*, 2016).

In this paper, we investigate the speech preferences of architects and engineers for conceptual CAD modeling using a speech- and gesture-based interface. The utility of this investigation is to examine whether speech- and gesture-based interfaces for conceptual CAD modeling for these two groups need to be differentiated based on professional affiliation. We use our findings to present recommendations for the design of a multi-modal CAD modeling interface.

Analytic themes for investigating verbalization

We investigate the speech preferences of architects and engineers for conceptual CAD modeling based on three analytic themes: (1) the exactness of the descriptions, (2) the granularity of the descriptions, and (3) the use of legacy knowledge.

(1) Precision is the exactness in design descriptions. We try to understand precision through the related concept of uncertainty, which is well researched in the design literature. Uncertainty is characterized by a lack of information, and includes vagueness and imprecision (Luck, 2013). Designers tend to employ imprecise, uncertain and provisional ideas in communication. Previous studies have reported that uncertainty is interwoven in design conversations (Luck, 2013). For instance, designers commonly employ vague expressions such as "*here*", "*this*", and "*there*" when speaking, often relying on gestures (Harrison and Minneman, 1996; Logan and Radcliffe, 2004). Such vague expressions are generally employed when the speaker does not have precise knowledge. Uncertainty in designers' verbalization is seen as appropriate for the early stages of design (Lawson and Loke, 1997).

Vagueness occurs whenever there is a need to specify structure, form or color approximately for later refinement (Fish, 2004). A number of terms such as "*ambiguous*" (Minneman and Harrison, 1998), "*vague*" (Harrison and Minneman, 1996), as well as "*fuzzy*" (Wiegers *et al.*, 2011) have been used in the design literature to refer to the uncertainty in design language used by designers. However, since in this study we are interested in the precision of descriptions, we refer to its opposing concept as vagueness and use it as an umbrella term to describe all imprecise or uncertain terms used by participants. Uncertainty and vagueness are known to characterize the conceptual design stage, as designers initially neither have complete information about the design problem, nor do they have clear ideas on how to address them. On the other hand, precision characterizes later stages of design, when designers specify details (Gross, 1996). Precision has been discussed previously in the context of CAD modeling (Walther *et al.*, 2007) and architectural design (Chastain *et al.*, 2002). The literature largely critiques extant CAD modeling systems for being overtly precision-based and argues that it does not suitably address the ways in which designers work in the conceptual design stage (Eckert *et al.*, 1999; Zheng *et al.*, 2001; Chastain *et al.*, 2002; Oh *et al.*, 2006; Zhong *et al.*, 2011).

The analytic theme of precision is especially important for conceptual CAD modeling using multi-modal interfaces, primarily for two reasons. First, due to the prevailing belief in current design research about conceptual design being characterized by vagueness (Lawson and Loke, 1997; Glock, 2009). Indeed, vagueness and ambiguity are considered important aspects of conceptual design (Goel, 1995), and are held significant for triggering reinterpretations (Eckert and Stacey, 2000). Second, there are technological issues in precise gesture recognition and command execution (Wang *et al.*, 2011). Wang *et al.* (2011) cite issues such as depth cues, selection of objects, and occlusions that encumber precise CAD modeling using gestural interaction. As a result, research in gesture-based interfaces for CAD modeling have largely focused on conceptual design, as the inaccuracy offered by gestural interaction is seen to be conducive for it (Alcaide-Marzal *et al.*, 2013). Therefore, it is important to investigate preferences of precision in user groups' natural articulation for conceptual CAD modeling using a multi-modal interface with speech and gestures.

(2) Granularity: Design representations are varied in their consistency. Depending on the purpose, design representations may provide a detailed account of all parts and aspects of the design artifact. At other times, design representations may be partial; they may pertain to certain elements only or they may display different components with varying amount of detail and attention (Herbert, 1988). We use the term "level of detail" to discuss granularity in participants' verbalizations. Whereas some representations are elaborate and detailed, others are rough outlines of initial ideas (Goldschmidt, 2004). Level of detail has been considered previously in the context of designers' speech (Logan and Radcliffe, 2004). In speech analysis, the level of detail pertains to whether designers choose to speak succinctly, or whether they provide detail about sizes, locations, and relationships. It is especially relevant to investigate the level of detail in designers' verbalizations for speech-based interfaces for CAD modeling, as it indicates how much detail designers prefer to incorporate in their instructions for conceptual design.

(3) Use of legacy knowledge: Legacy knowledge is based on users' experience with prior interfaces and technologies (Morris *et al.*, 2014; Beşevli *et al.*, 2018). Research in elicitation studies has found that previous experience with desktop computing strongly influences users' gestural responses (Morris, 2012). Since the current generation of architects and engineers are well versed with CAD modeling software that relies on WIMP, it may be assumed that when speech is elicited from them, their responses would be affected by their knowledge, experience, and habit of working with WIMP-based interfaces. Previous approaches in gesture elicitation fall under two categories: (1) studies that aim to reduce legacy bias (Morris *et al.*, 2014) and (2) studies that aim to benefit from it (Köpsel and Bubalo, 2015). The former approach argues that legacy bias limits the potential of user-elicitation

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

> methodologies for producing interactions that take full advantage of emerging application domains. On the other hand, the latter approach reasons that legacy knowledge would make it easier to design new interactions, as familiar knowledge is easier to recall, produces confidence and is especially useful for specific user groups (Köpsel and Bubalo, 2015).

In our study, we use the phrase "CAD legacy terms" to refer to terms that are used in established CAD programs such as AutoCAD, SketchUp, 3dsMax, and Solidworks. Our interest in the use of legacy knowledge is to investigate the kind of terms designers would employ; and if there is any difference in the terms employed by the designers, based on their professional affiliation. Do any professional groups use certain CAD modeling terms more than others?

Investigation in the language verbalized by designers is indispensable, as words are an integral part of the design process in the early stages of design. Previous studies have largely employed protocol analysis techniques for studying the language employed by designers (Athavankar, 1999). Other studies have employed natural language processing techniques to study design communication (Dong, 2005). Research into the compilation of terms for design includes a study by Podehl (2002) on styling terms. A noteworthy study by Wiegers *et al.* (2011) compiled the terms for shapes and operations employed by designers. Cicognani and Maher (1997) presented a list of verbs for use in virtual communities for design. Investigation into the ambiguity and uncertainty in design communication includes ethnomethodological approaches using conversation analysis (Luck, 2013).

Although extant literature lays importance on the investigation of linguistics for system design (Luck, 2013), so far there exist little research into the words that designers use for CAD modeling for conceptual design. We assert that speech-based human–computer interfaces must employ speech terms that are natural for specific user groups, adapting to the language that is commonly employed by them. Hence, we propose that the vocabulary set of speech- and gesture-based CAD modeling interfaces must be informed by user behavior. We addressed this issue by analyzing CAD modeling terms extracted from an experiment with participants from architecture and engineering backgrounds, elaborated in the following section.

Method

Participants

The study presented in this paper is based on data collected from a gesture and speech elicitation experiment. As described previously in our studies (Khan *et al.*, 2017; Khan and Tuncer, 2017; Tunçer and Khan, 2018), the experiment was conducted individually with 20 engineers from an EPD background and 20 architects. EPD is a combination of the traditional disciplines of mechanical engineering and electronics and electrical engineering. The product sectors that are primarily addressed in EPD are electronics, energy, machinery, and transportation. Out of the 40 participants, 21 were female, and 19 were male. The experiment was conducted over a period of 2 weeks at the Singapore University of Technology and Design. Participants comprised the following ethnicities: Chinese (57.5%), Indian (22.5%), Caucasian (10%), and other (10%). Participants consisted of undergraduate students (25%), Masters students (20%), PhD students (10%), researchers (17.5%), faculty members (17.5%), and practitioners (10%). Most of the participants were in the 22–30 years age group (65%), followed by 31–40 years age group (15%), and the 18–21 years age group (10%). Although 80% of the participants reported English as their first language, all participants were fluent in English, which was a prerequisite for participation in the experiment. More than 90% of the participants reported being acquainted with one or more CAD software.

The sample size was the standard used in speech and gesture elicitation studies, as evidenced by previous studies that have employed similar sample sizes ranging from 20 to 30 participants (Wobbrock *et al.*, 2009; Morris, 2012). In studies that investigate differences between architects and laypersons, Gifford *et al.* employed a sample size of 17 architects (Gifford *et al.*, 2002).

Experimental design

The aim of the experiment was to elicit speech and gestures that communicated CAD modeling tasks such as creating and manipulating 3D objects and navigating views. The participants sat at a distance of 10′ from a 50″ sized screen where a repertoire of prerecorded CAD modeling tasks (referents) were shown one by one, in the form of short video clips. The participants were asked to describe the CAD modeling tasks shown on the screen. The categories and referents were randomized for all participants.

The referent tasks were all low level, basic CAD modeling operations classified into three categories: (1) Navigation, which involved changing the view, (2) Manipulations, and (3) Primitives. In this study, we investigate participants' verbalizations for four basic manipulation tasks in CAD modeling, namely Scale, Rotate, Move, and Copy (Fig. 1).

The object of manipulation in the first three referent tasks was a basic box. For Scale, a video clip was shown with two boxes in the first frame. When the video clip was played, the bigger box scaled up uniformly in all directions. In the video clip for Rotate, a box rotated 45° toward the right-hand side. For Move, two boxes were shown in the first frame of a video. When the video clip played, the shorter box slowly moved along the x-axis. For Copy, a compound object was shown in the first frame. When the video clip played, the object duplicated toward the right-hand side. Each video clip was shown for approximately 15 s.

The experiment was conducted in two sessions, A and B. In the pre-test briefing, participants were given a scenario in which they were informed that Laura was a designer sitting in the other room and needed assistance in manipulating the object as they see in the video clips on their screens. In session A, participants were informed that Laura could only see them and not hear them. Therefore, in session A, participants articulated the referent tasks using only gestures. In session B, participants were told that Laura could see them as well as hear them. Therefore, in session B, participants were free to use hand gestures or speech. There was no restriction on the length or the technique of the instruction. Participants employed spontaneous, free speech. The sessions were video recorded using two high-speed cameras from different angles. The participants were queried about their educational, professional, and socio-cultural background in a questionnaire. In this paper, we report findings from session B of the experiment.

Coding

The video recorded data were edited into named clips. We transcribed 160 records (2 groups × 20 participants × 4 manipulation

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

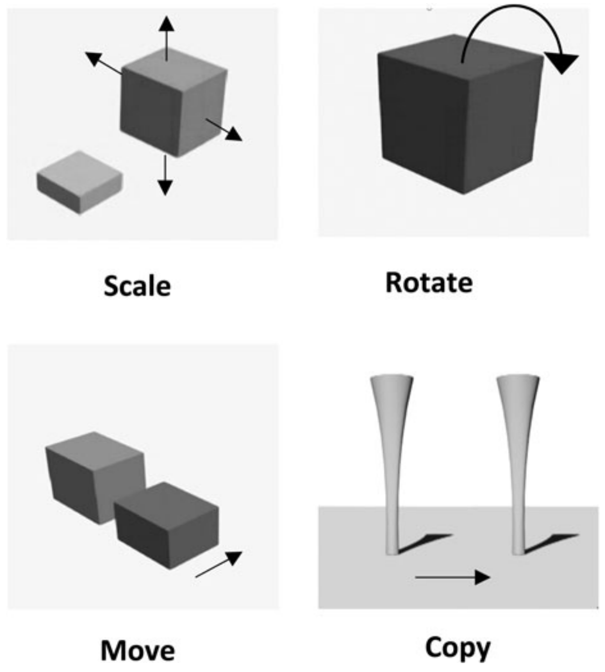

Fig. 1. Manipulation tasks used in the experiment.

Table 1. Coding categories for each manipulation task

Task	Coding categories
Scale	1. Command 2. Object 3. Dim: degree of scale 4. Dim: aspect of the object being scaled 5. Location
Rotate	1. Command 2. Object 3. Dim: angle of rotation 4. Axis of rotation 5. Direction of rotation
Move	1. Command 2. Object 3. Original location 4. Direction of movement 5. Dim: Distance
Copy	1. Command term 1 2. Command term 2 3. Object 4. Target location 5. Dimension: target number of copies

tasks) of participants' speech for manipulation tasks. A hybrid coding approach was followed to decide the categorization scheme. Transcription scripts were first reviewed jointly by both co-authors and a PhD researcher to identify the initial set of categories. Thereafter, a single coder (co-author of the paper) carried out the coding, with periodic joint reviews to resolve differences. The categorization system was further improved as the coding proceeded.

In the coding, we focused on the part where participants described the object to manipulate and how to manipulate it. Parts where they digressed were ignored. We ignored the tenses of verbs and used their infinitive form. For instance, "*moving*" or "*move*" or "*moved*" were all categorized as one. We considered all synonyms and did not generalize them. Repetitions and erroneous usage of words were ignored.

Based on the function performed in the verbalizations, we extracted words from the speech transcriptions of the participants and classified them into the following categories:

- *Command*: words that instruct to execute the manipulation.
- *Object*: words that describe the object being manipulated.
- *Dimensions*: expressions that indicate units or the degree of manipulation, such as angle of rotation, distance, and the number of copies.
- *Location*: expressions that indicate the
  - Original location or position of the object
  - Target location or position of the object
  - Directions, or relative position of a target from a point of origin.
- *Dimensional aspects*: examples include side, size, and volume
- *Modifier*: conditional words that restrict or modify the command.

We listed the categories relevant for each manipulation task (Table 1). Based on these categories, each verbalization was coded. Table 2 shows an example of coding a verbalization for the modeling task Scale. The frequency distributions of all categories were determined for architects and engineers for each manipulation task. We developed measures for determining the precision and level of detail in participants' verbalizations.

Key definitions

CAD legacy and non-legacy command terms

We defined CAD legacy command terms based on the terms used for manipulation tasks in the CAD programs used by the participants. These were, namely, AutoCAD, Rhino, 3dsMax, SketchUp, and Solidworks. Terms classified as CAD legacy commands are given in Table 3. Remaining command terms employed by participants, that are not used in the aforementioned CAD software, were classified as non-legacy.

Precision score

Precision scores were calculated based on the number of precise and vague terms present in a given verbalization. Examples of such precise and vague terms used by the participants in the experiment are presented in Table 4. Every precise term in a verbalization was given a positive point, whereas every vague term was given a negative point. Thus, the precision score for a verbalization was calculated as:

$$P = n_1 - n_2$$

where n_1 is the number of precise terms and n_2 is the number of vague terms.

Therefore, if the number of precise terms in a verbalization was greater than the number of vague terms, the precision score was positive. If the opposite was true, the score was negative. If there were an equal number of precise and vague terms in a verbalization, the score was zero.

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

Table 2. Example of coding a verbalization for the modeling task Scale

Verbalization	Command	Object	Dim: degree of scale	Dim: aspect of the object being scaled	Location
Extend the block to about three times its size around its *x*-axis	Extend	Block	3 times	Size	*x*-axis

Table 3. CAD legacy command terms from the existing CAD software

	AutoCAD, Rhino	3dsMax	SketchUP	Solidworks	CAD legacy command terms
Scale	*scale*	*scale*	*scale*	*scale*	*scale*
Rotate	*rotate*	*rotate*	*rotate*	*rotate*	*rotate*
Move	*move*	*move*	*move*	*move*, *translate*	*move*, *translate*
Copy	*copy*	*clone*, *copy*, *instance*	*copy*	*copy*	*copy*, *clone*, *instance*

Table 4. Examples of precise and vague terms used by participants in the experiment

		Dimensions		Locations		
	Vague modifiers	Precise	Vague	Precise	Vague	Vague expressions
Scale	*about*; *maybe*; *slightly*; *a bit*; *or so*; *around*; *approximately*; *a little more than*; *a little less than*; *probably*;	*x times*; *double*; *twice* *x %*	*bigger*; *larger*; *larger and larger*; *this big*	*x*, *y*, *z axis*	*equally*; *uniformly*	*a bit like this*; *like this*; *this way*; *like that*; *this position*; *kind of like this*;
Rotate		*x degrees*; *x times*	*this angle*; *like this*; *an angle*;	*Clockwise* *downward* *x or y direction*	*about its edge*; *about its vertex*	
Move		*x units*; *x times*; *double*; *two-thirds*; *x %*	*until this point*; *this distance*; *small distance*;	*left*; *right*; *forward*;	*away*; *nearer*, *farther*; *to one side*; *the other way*; *sideways*;	
Copy		*x*, *y coordinates*	*this distance*; *this angle*; *small distance*	*right*; *horizontally*; *linearly*	*slightly behind*; *to its side*; *next to it*;	

Detail score

Level of detail score (D) was computed based on the number of coding categories a participant used to verbalize a given task (Table 1). For every coding category used in the verbalization, a participant was assigned a value of 1. We defined the level of detail as the sum of individual coding categories present in the description, on a five-point scale. Accordingly, the score was greater with the increase in the number of categories used in verbalizations:

$$D = \sum_{i=1}^{5} p_i$$

where p_i is a coding category which can take the value of 0 or 1.

Results

A total of 3404 words were transcribed from the participants' verbalizations of the four manipulation tasks. The greatest number of words were used by the architects' group to verbalize the Move task ($Mdn = 22.5$ words), and the least number of words were used by the architects' group to verbalize the Copy task ($Mdn = 14$ words) (Fig. 2). For the Rotate and Move tasks, the median number of words spoken by the architects' group was greater than the median number of words spoken by the Engineers' group. On the other hand, for the Scale and Copy tasks, the median number of words spoken by Engineers was greater than that of Architects.

Compilation and legacy terms

We compiled the terms used by the participants for articulating the CAD manipulation commands (Fig. 3). The task Scale had the greatest diversity in command terms, with five different command terms used by at least 5% of the participants. The task Rotate had the least diversity in command terms with only two different command terms that were used by at least 5% of the participants or more.

We investigated the use of CAD legacy command terms (Table 3) in participants' verbalizations. Overall, CAD legacy terms were used by a majority of the participants (over 60%) in

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

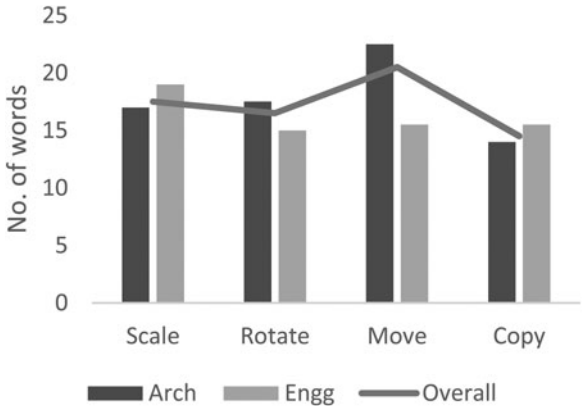

Fig. 2. Median verbalization length of architects and engineers for manipulation tasks.

the case of Rotate and Move. For the tasks Scale and Copy, the overall use of CAD legacy terms was a little over 40%.

More than 70% of the architects employed the CAD legacy terms in the case of Scale and Rotate tasks, whereas their use of legacy terms was close to 60% in the case of Move and Copy. A χ^2 test also revealed a statistically significant difference between architects' and engineers' use of legacy and non-legacy terms for the tasks Scale and Rotate (Table 5). The majority of Engineers employed CAD legacy terms only in the case of Move. For the other three manipulation tasks, the use of CAD legacy terms by engineers was 50% or less.

Precision scores

We studied median precision scores for both professional groups. We found that except for the median score of Copy for architects ($Mdn = 1$), the median score for other categories for both architects and engineers was 0, implying that most participants used an equal number of precise and vague expressions in their verbalizations, or had verbalizations that scored midway between precise and vague. Examples of such verbalizations with a precision score of 0 include:

> "*Enlarge.*" (Scale task, Participant 24, Engineer)
> "*There is a cube which is in this angle and just rotates and comes to this position.*" (Rotate task, Participant 39, Engineer)
> "*Move the block on the right, away.*" (Move task, Participant 2, Architect)
> "*There is one item, copy one more.*" (Copy task, Participant 5, Engineer)

A majority of architects (55%) employed precise language only in the case of Copy. An example of a verbalization that was scored as precise was

> *The object is front of you, I want to make a copy of it and move it along the ground 1.5 times its width.* (Copy task, Participant 16, Architect)

In all other cases, the numbers were somewhat evenly distributed across the positive, neutral, and negative categories in the histogram (Fig. 4). Comparison of the frequency chart of the precision scores showed that a slightly greater number of architects used precise language than engineers in the cases of Scale (Arch = 35%, Engg = 15%), Rotate (Arch = 40%, Engg = 30%), and Copy (Arch = 55%, Engg = 35%). However, results from the Mann Whitney U test indicated that there was no significant statistical difference between the precision scores of architects and engineers for all four manipulation tasks (Table 6).

We investigated the cases in which participants employed precise dimensions vs. those in which participants gave vague dimensions (Fig. 5). Precise dimensions involved giving exact numbers in units such as degrees or percentage. In the case of vague dimensions, participants employed language such as "*about this much*" or "*from here to there*", using gestural cues (Table 2). Overall, a greater percentage of participants gave precise dimensions for the tasks Scale and Rotate. On the other hand, most participants gave no dimensions for the tasks Copy and Move. Close to 45% of Architects employed precise dimensions for the tasks Scale and Rotate. Except for the Rotate task, for which close to 45% of engineers employed precise dimensions, most engineers did not give any dimensions for the other three tasks.

Detail scores

We counted the number of coding categories participants used for verbalizing tasks. The median D scores for all four manipulation tasks for both architects and engineers was 3, except for the median score for architects for the task Rotate, which was 3.5. This implies that participants used a median of three parameters in their verbalizations, primarily specifying the manipulation command and the object to manipulate, with one other variable parameter, such as the direction, dimension, or another aspect of manipulation. Following are examples of verbalizations with a level of detail score of 3:

> "*Increase the size of the bigger object.*" (Scale task, Participant 3, Engineer)
> "*Rotate the object to its side.*" (Rotate task, Participant 20, Engineer)
> "*Push the shorter block like this.*" (Move task, Participant 35, Architect)
> "*You have an object and you clone it to the right.*" (Copy task, Participant 19, Architect)

A comparison of the means for each manipulation task shows that the mean for Architects was greater than the mean for Engineers by a very slight margin (Fig. 6). Results from the Mann Whitney U test indicated that there was no statistical difference between the level of detail scores of architects and engineers for all four manipulation tasks (Table 7).

Discussion

The goal of this study was to investigate the speech preferences of architects and engineers for conceptual CAD modeling using speech- and gesture-based interfaces. We sought to investigate questions such as

- Do both user groups prefer to use precise expressions, or do they use vague expressions?
- Do engineers employ more legacy terms than architects?
- Is the language employed by these groups influenced by their knowledge of existing CAD software?

We investigated the speech terms employed by participants from the two professional groups for four CAD manipulation tasks. We examined the verbalizations based on their exactness, granularity, and use of legacy knowledge.

We thus further the intent of previous studies (Wiegers *et al.*, 2011) in presenting an analysis of designers' verbalizations of manipulating objects, specifically for the purpose of the design

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

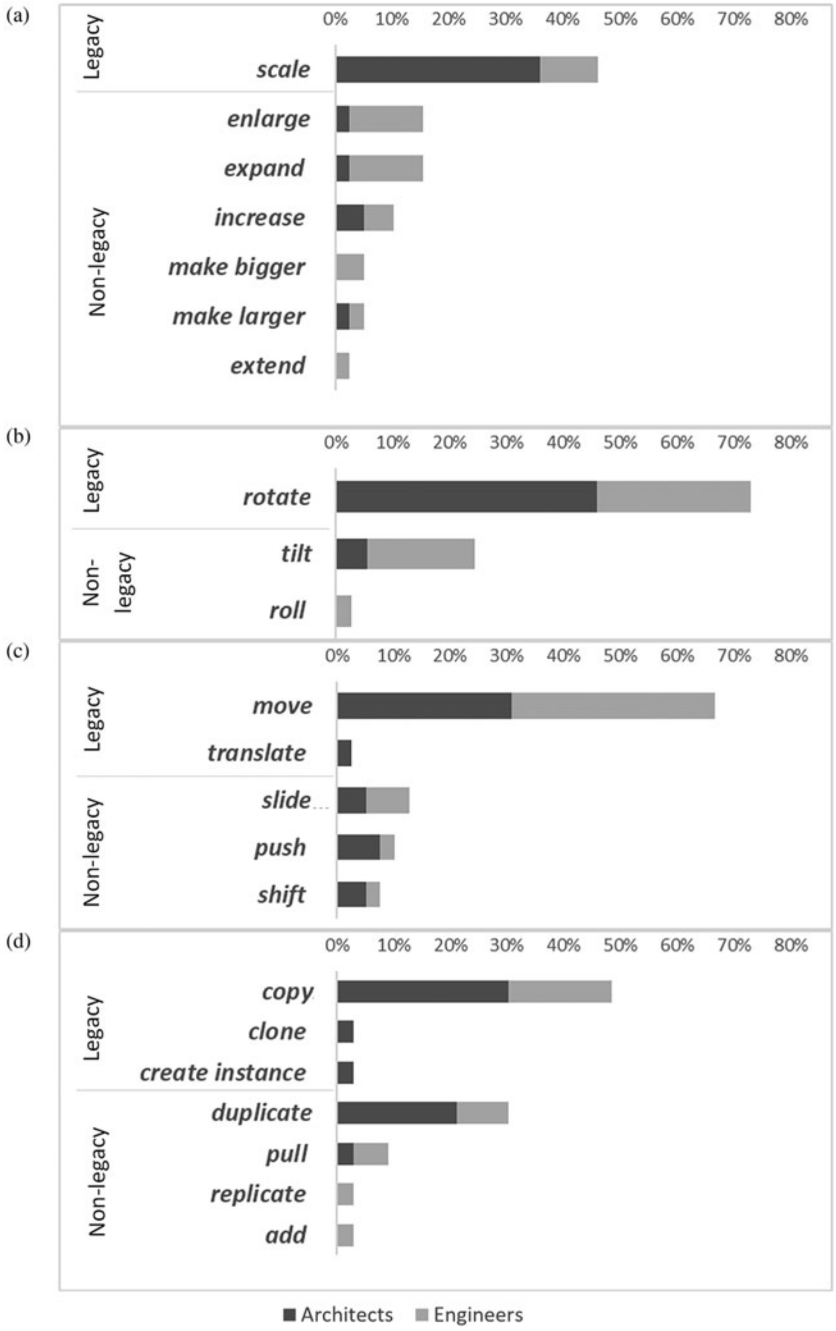

Fig. 3. Speech terms used for manipulation tasks: (a) Scale, (b) Rotate, (c) Move, and (d) Copy.

of an interface for speech- and gesture-based conceptual CAD modeling. In this section, we discuss our findings and based on these, present recommendations for the design of a conceptual CAD modeling interface using speech and gestures.

Exactness of descriptions

We investigated the use of precise and vague expressions by architects and engineers. We found that for most manipulation tasks, a comparable number of participants employed precise expressions and vague expressions in their verbalizations. For example, when describing dimensions, a greater percentage of participants gave precise dimensions in units or degrees for the tasks Scale and Rotate, while most of the participants gave no dimensions for the tasks Copy or Move. This could be attributed to the greater geometrical complexity of the Scale and Rotate tasks. In the use of precise dimensions, a number of participants added vague modifiers, using phrases such as "*maybe* 10%–20%" and "*slightly more than* 45 degrees". This could be attributed to the increased physical distance between participants and the screen when using gestural interaction, which led them to approximate distances when using precise units. Thus, we deduce that designers from the two professional groups use both precise and vague expressions for manipulation in conceptual design, depending on the context and the nature of the task. This crucial finding is in direct contrast to previous studies that have concluded that designers prefer to employ imprecise vocabulary in their speech rather than select words with precise meaning (Logan and Radcliffe, 2004).

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

Table 5. Results of the χ^2 test: difference between architects' and engineers' use of legacy and non-legacy terms ($N = 40$)

	χ^2 statistic	*P*-value	
Scale	10.101	0.001482	Significant
Rotate	5.584	0.018125	Significant
Move	0.114	0.73568	Not significant
Copy	3.6364	0.05653	Not significant

$^*p < 0.05$.

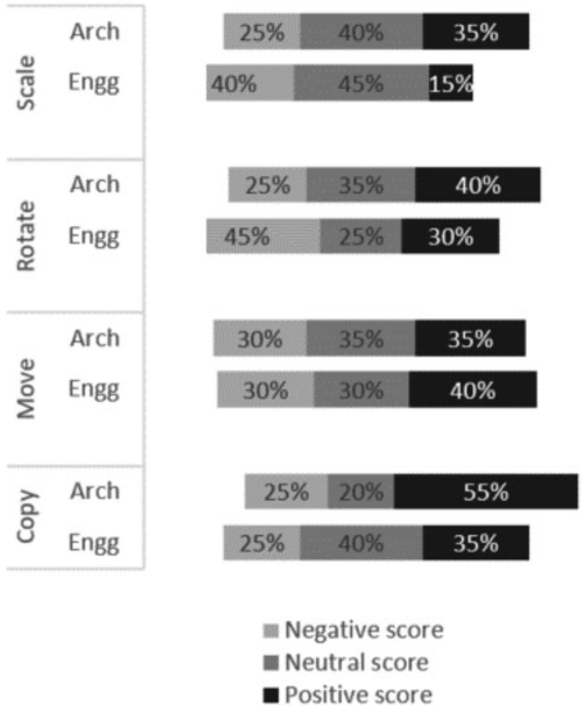

Fig. 4. Frequency distribution of precision scores for architects and engineers (positive scores indicate a greater number of precise expressions in verbalizations, negative scores indicate a greater number of vague expressions).

Table 6. Precision scores: results from the Mann Whitney *U* test

	U-Value	*Z*-Score	*p*-Value
Scale	154	1.23	0.218
Rotate	154.5	1.21	0.222
Move	198	−0.040	0.968
Copy	180	0.527	0.596

$^*p < 0.05$.

This has a very significant implication on the design of speech- and gesture-based interface for conceptual CAD modeling. This suggests that at times, depending on the context, designers want to give precise dimensions. At other times, they want to employ vague instructions using gestural cues and speech. Therefore, we recommend that a speech- and gesture-based interface for conceptual CAD modeling should recognize and support such a precision-vagueness dichotomy and allow designers to switch from one mode to another. For example, if a designer gives a vague instruction "Move box a little bit (using gestural cues)", the interface should allow the gestural instruction to override the speech instruction. On the other hand, if the designer gives a more precise instruction such as "Move box 5 inches to the left (using gestural cues)", the interface should let the speech input override the gestural input.

Our recommendation for the support of precision-vagueness dichotomy is also substantiated by previous literature that argues for the development of interfaces that mimic the way people naturally communicate (Cassell, 1998; Quek *et al.*, 2002). Although the precision-vagueness dichotomy may seem obvious in the context of natural speech, our recommendation is significant in the context of conceptual design, in which uncertainty and ambiguity are seen to prevail in the current literature (Lawson and Loke, 1997; Glock, 2009), and precision-based CAD modeling systems are critiqued for being unsuitable for conceptual design (Eckert *et al.*, 1999; Zheng *et al.*, 2001; Chastain *et al.*, 2002; Oh *et al.*, 2006; Zhong *et al.*, 2011). While Stacey and Eckert (2003) present an argument against ambiguity for communications in conceptual design, we provide empirical evidence to show that architecture and engineering professionals do not always use only vague language to communicate conceptual CAD manipulation tasks. As opposed to interfaces that make users learn an artificial, author-defined vocabulary, support of precision-vagueness dichotomy in an interface would give flexibility and choice to users.

Although a slightly greater number of architects than engineers used precise expressions, there was no statistical difference between the precision scores of the two groups. Thus, we deduce that for basic manipulation tasks in CAD modeling, professional affiliation had little or no significant bearing on the language employed by the participants. This finding is in direct contrast to previous studies that conclude that education has a bearing on the way subjects communicate shapes and shape operations (Wiegers *et al.*, 2011). We attribute our contrasting finding to the basic, low-level nature of the given manipulation tasks. The tendency of architects to use precise expressions is an avenue that ought to be explored further with more experiments with conceptual CAD modeling tasks of greater complexity.

Compilation and legacy knowledge

Based on analyses of participants' verbalizations, we compiled a set of command terms from the experiment that can be used for conceptual CAD modeling in multi-modal interfaces using gestures and speech.

A noteworthy finding was the use of non-legacy terms by 35%–45% of the participants in all cases, even though more than 90% of our participants reported being acquainted with one or more CAD software programs. Even though two of the terms classified as legacy, "Copy" and "Move", are also employed in everyday usage, we found nonetheless that in all four cases, around 35%–45% of the participants used non-legacy command terms for the description of the manipulation tasks. We deem that users employ a range of words to describe manipulation tasks in their day to day communication, and hence legacy terms should not be forced on all users.

We found that Architects were well versed with CAD legacy terms and employed them in more cases than the Engineers. This may be attributed to architects having greater working knowledge and experience with CAD software than engineers. It also suggests that the current generation of architects, who have

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

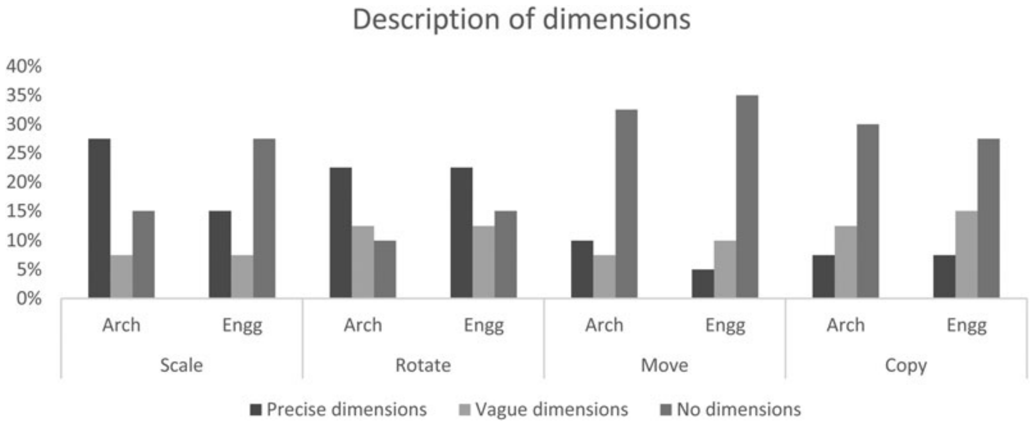

Fig. 5. Description of dimensions by the two professional groups for the four manipulation tasks.

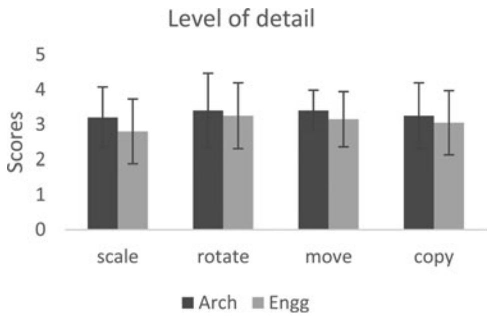

Fig. 6. Level of detail: mean scores for architects and engineers.

Table 7. Level of detail: results from the Mann Whitney *U* test

	U-Value	*Z*-Score	*p*-Value
Scale	155.5	1.190	0.234
Rotate	179.5	0.541	0.541
Move	169.5	0.811	0.811
Copy	172.5	0.730	0.465

$^*p < 0.05$.

been trained in CAD in university or soon after, will initially rely on their conceptual knowledge of existing CAD software for CAD modeling, even though natural interfaces offer vastly different interaction techniques. This could especially be true for more geometrically complex commands, as evidenced by Architects' greater use of legacy terms for the tasks Scale and Rotate.

Hence, as opposed to previous studies that seek to reduce legacy bias in human–computer interaction (Morris *et al.*, 2014), we take an inclusive approach and view the knowledge of existing CAD terminology as also relevant. We, therefore, recommend the inclusion of legacy terms as well as non-legacy terms. We reason that present-day design professionals are trained in CAD in their early years of education and are hence well versed with CAD terminology. Therefore, we find little reason to discard this collective, accumulated knowledge of legacy terms. Such an approach also helps shorten the time and effort required by professionals to learn the new ways of interaction (Köpsel and Bubalo, 2015). Our recommendation for the inclusion of both legacy and non-legacy knowledge in the interface is supported by evidence from recent gesture elicitation research, which found that legacy gestures were favored by participants for their familiarity and non-legacy gestures were favored for their affordances (Beşevli *et al.*, 2018). This suggests that both legacy and non-legacy knowledge is useful, based on user needs and context.

Therefore, our compilation includes all viable terms that were employed by the participants in the user experiment, as listed in Figure 3. We also recommend a many-to-one mapping of speech command terms to CAD functionalities, for the design of a speech- and gesture-based CAD modeling system. Based on an initial set, a user should be able to modify or extend the speech set in the system or the system should incorporate machine learning to adapt to the user. We assert that this approach is more aligned with natural human interaction and will result in the development of an interface that is flexible and attuned to the needs of different users. Previous studies that use a similar approach include the studies of Coroado *et al.* (2015) in which commands follow a context-free grammar and each operation can be triggered by one or more voice commands. A similar approach of using synonyms is suggested in gesture elicitation studies by Wobbrock *et al.* (Wobbrock *et al.*, 2009), to increase guessability and the coverage of proposed gestures. For multimodal interactions, Morris (2012) suggested the use of multimodal synonyms which would allow users to access the same functionality with different modalities in different circumstances. As opposed to previous studies that employ arbitrary or author-defined speech commands (Nanjundaswamy *et al.*, 2013), our compilation is more thorough and informed by user behavior.

Granularity of descriptions

An investigation into the median level of detail scores of Architects and Engineers revealed that most participants from both professional groups employed three parameters to communicate the manipulation tasks. The three parameters described the manipulation command, the object to manipulate and how to manipulate it.

Therefore, we recommend that a speech- and gesture-based interface for conceptual CAD modeling should allow users to verbally describe three parameters for manipulation tasks, based

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

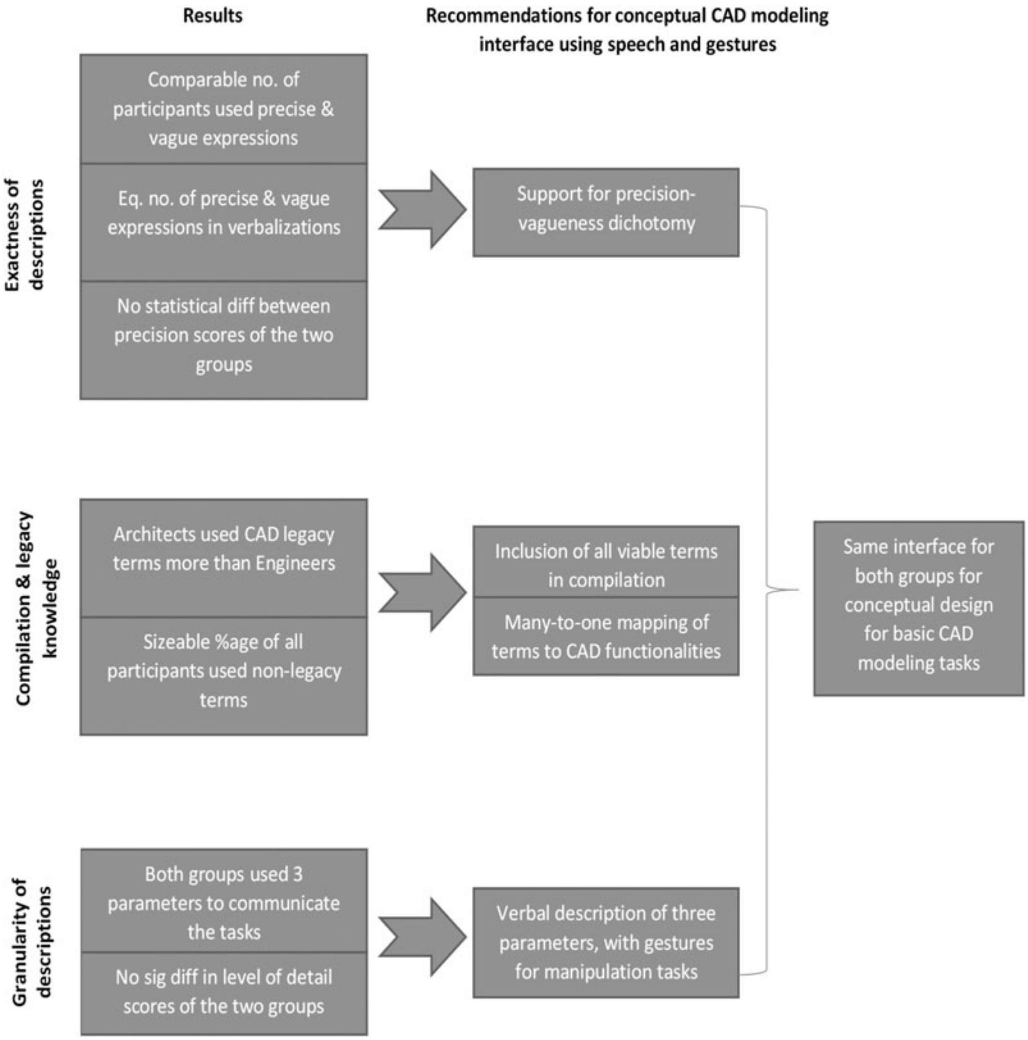

Fig. 7. Summary of results and recommendations for the design of a conceptual CAD modeling interface using speech and gestures.

on users' natural articulation. While the command term and object name are imperative for invoking the CAD functionality, a third variable parameter, such as direction, dimensions, or another aspect of manipulation, along with gestural cues, should give sufficient information to the system to carry out the manipulation task. This recommendation builds upon previous studies in multi-modal interfaces that only employ single word commands for CAD modeling (Nanjundaswamy *et al*., 2013). Our recommendation is also supported by the literature on multi-modal interfaces that argue for the use of short speech strings (Varshney, 1998) and the use of brief, syntactically simpler language (Oviatt, 1999).

Although previous studies have found that architects tend to give slightly more detail than non-architects (Devlin, 1990), our analysis did not show a significant difference in the mean level of detail scores of architects and engineers. This is also attributable to the basic nature and low semantic level of the referent manipulation tasks shown to the participants, and we deduce that the difference in scores for the two professional groups would be significant for more complex referents such as buildings. This finding is also aligned with the results from Logan and Radcliffe (2004) who concluded from their study that designers speak with a simple vocabulary in their speech and employed a visual channel to aid in the verbalizations.

Conclusion

In this paper, we presented an analysis of verbalizations for conceptual CAD modeling extracted from a specially conducted experiment with architects and engineers. We compiled the command terms that the two professional groups employed to describe four basic CAD manipulation tasks. We presented insights into the choice of command terms, and preferences of precision and detail in the verbalizations of architects and engineers.

Summary: chief findings and recommendations

We summarize here our chief findings and the recommendations for a multi-modal interface for conceptual CAD modeling (Fig. 7):

- We found that a comparable number of participants used precise and vague expressions in their verbalizations, and that most

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

participants employed an equal number of precise and vague expressions in their verbalizations. We deduced that designers from the two professional groups use both precise and vague expressions for manipulation in conceptual design, depending on the context and the nature of the task. Therefore, we recommend that a multi-modal interface for conceptual design must support a precision-vagueness dichotomy.

- We found that Architects used CAD legacy terms more than engineers in two of the four manipulation tasks; and that a sizeable percentage of participants used non-legacy terms. Therefore, we recommend the inclusion of all viable terms in the compilation and a many-to-one mapping of terms to CAD functionalities.
- We found that participants from both groups used a median of three parameters to articulate the tasks: the manipulation command term, the object to manipulate and a variable parameter such as direction, dimension, or another aspect of manipulation. Therefore, we recommend the verbal description of three parameters along with gestures for manipulation tasks.
- We found no statistical difference in precision scores and level of detail scores of the two groups. We deduced that for basic manipulation tasks in CAD modeling, professional affiliation had little or no significant bearing on the language employed by the participants. Therefore, we recommend that the same interface for conceptual design could be used by both architects and engineers, for basic CAD modeling tasks.

We highlight that our recommendations are based on a user-centered approach, whereas in extant studies natural interfaces are often critiqued for not being natural (Malizia and Bellucci, 2012). As opposed to previous studies in multi-modal interfaces that provide a predefined speech set with a one-to-one mapping of words to CAD tasks (Nanjundaswamy *et al.*, 2013), our recommendations reflect a soft approach that allows users to choose from a number of verbal terms to initiate commands. We assert this approach to be closer to natural human communication, as it gives users flexibility and choice (Cassell, 1998). Furthermore, our recommendations for multi-modal interfaces are based on empirical evidence and substantiated with the literature from human–computer interaction and natural interfaces. Future research on multi-modal interfaces that employ speech could immensely benefit from our user-centered recommendations. We conducted a preliminary testing of our recommendations and note some of the strengths and challenges in the implementation of a prototype (Khan *et al.*, 2017).

Undoubtedly, the bigger challenge here is the development of a robust recognition system that can incorporate complexities such as multiple words for same CAD functionality, and speech-gesture overrides, which are routine functions in human to human interaction. Taken forward, artificial intelligence techniques are indispensable for the implementation of our recommendations, as such complexities can be suitably addressed by natural language processing and learning algorithms.

Limitations and future work

We analyzed verbalizations by developing measures to count the number of expressions in participants' verbalizations. As opposed to the approaches of protocol analysis (Gabriel and Maher, 2002) and conversation analysis (Luck, 2013) that have been used in previous studies of designers' communication, we employed a content analysis approach of coding and counting expressions in verbalizations to investigate their tendencies of precision and detail. We conceptualized precision and vagueness as bipolar opposites on a balanced scale from −2 to 2; and level of detail as a positive scale from 0 to 5. Given the basic nature of the CAD tasks, and hence the short length of verbalizations, these five-point scales were considered sufficient for this study. A similar approach to analyze the level of detail is found in the study by Logan and Radcliffe (2004), who analyze details in designers' verbalizations by counting the number of nouns, verb, and adverbs. Such an approach is seen to be appropriate as a simple test of a verbalization's descriptive specificity (Logan and Radcliffe, 2004). We acknowledge that more complex measures can be developed, for instance, by adding weights to the counts of expressions, based on a given rationale.

Categorization of terms as legacy or non-legacy knowledge can be contentious. For instance, words that are employed in day to day language as well as in CAD software, such as "*move*" or "*copy*" – are these CAD legacy terms or non-legacy? Our categorization of legacy terms was based on two criteria: (1) background of the participants, and (2) the context of usage of the terms. Since more than 90% of the participants reported to be acquainted with one or more CAD software, and the context of the verbalization was also CAD-based (manipulation of CAD objects on a plain background), we categorized these terms as legacy. We reason that had the participants been shown a different context, for instance, water moving in a river or cars in the traffic – in that case, it would have been more rational to categorize the word "*move*" as non-legacy.

Spoken language depends on the socio-cultural, educational, and linguistic background of people. We acknowledge that our study is tilted toward the accepted linguistic norms of the English-speaking populace of Singapore. The issue of a first language or mother tongue is complex in Singapore, as most people born after 1970s are considered bilingual, with an equal amount of proficiency in English as the language considered to be the first language or mother tongue. English is the medium of instruction in educational institutes, as well as the lingua franca for professional practice (Tan, 2014). Therefore, that only 80% of participants reported English as their first language should not be a cause for concern for this study.

Our focus on the three analytic themes was due to their relevance to conceptual CAD modeling using speech: how much to speak, how precisely to speak and what kind of terms to use. We consider our study as an initial investigation and acknowledge that other aspects of speech, for instance, fluency and errors (Oviatt *et al.*, 1997) in the context of CAD modeling, are also relevant and worthy of investigation in future studies. Furthermore, this research investigated the language designers use for basic manipulations. Conceptual design often involves complex operations performed on various kinds of geometry. Greater investigation is required into the language designers would use to describe complex operations such as Boolean operations, extrusion, and irregular shapes. A robust multi-modal interface requires strategic integration and synchronization of different modes in the system (Oviatt, 1999). Hence, a truly natural interaction system for CAD modeling would recognize the interplay of speech and gesture, its nuances and vague expressions, as envisaged in Bolt's "put-that-there" system (Bolt, 1980).

Uncertainty in the creation and manipulation of objects in conceptual design is fundamentally different from the precision-based input typical of extant CAD software. This research demonstrated how designers used a combination of precise and vague

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

expressions in their verbalizations. Conceptual CAD modeling requires the development of software, which not only supports such needs of dichotomy but possibly uses it to an advantage. Human communication is undoubtedly complex and nuanced, and an interface that seeks to incorporate the nuances of natural human interaction will need to address the challenges and not circumvent them.

This research provided insights into how architects and engineers convey information about conceptual CAD modeling tasks through speech. Such investigations are necessary for the successful design of a speech- and gesture-based conceptual CAD modeling interface. The strength of our approach is that it builds on the language that designers naturally employ and is informed by user behavior. Due to the complexities of natural language interpretation in conjunction with gesture recognition, a successful implementation of such an approach firmly relies on artificial intelligence techniques such as natural language processing and learning.

Financial support. This research was supported by SUTD-MIT International Design Centre (IDC) grant number IDG21500109, under the Sustainable Built Environment Grand Challenge, and Visualization and Prototyping Design Research Thrust.

References

Akalin A, Yildirim K, Wilson C and Kilicoglu O (2009) Architecture and engineering students' evaluations of house façades: preference, complexity and impressiveness. *Journal of Environmental Psychology* **29**, 124–132.

Akin O and Moustapha H (2004) Strategic use of representation in architectural massing. *Design Studies* **25**, 31–50.

Alcaide-Marzal J, Diego-Mas JA, Asensio-Cuesta S and Piqueras-Fiszman B (2013) An exploratory study on the use of digital sculpting in conceptual product design. *Design Studies* **34**, 264–284.

Athavankar U (1999) Gestures, mental imagery and spatial reasoning. In Gero JS and Tversky B (eds), *Preprints of the International Conference on Visual Reasoning (VR 99)*. MIT, pp. 103–128.

Berlyne DE (1971) *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.

Beşevli C, Buruk OT, Erkaya M and Özcan O (2018) Investigating the Effects of Legacy Bias. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility – DIS '18* (pp. 277–281). New York, New York, USA: ACM Press.

Bolt RA (1980) "Put-that-there": Voice and Gesture at the Graphics Interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques – SIGGRAPH '80*. (vol. **14**, pp. 262–270). New York, New York, USA: ACM Press.

Brown DC, Kwasny SC, Chandrasekaran B and Sondheimer NK (1979) An experimental graphics system with natural language input. *Computers & Graphics* **4**, 13–22.

Cassell J (1998) A framework for gesture generation and interpretation. In Cipolla R and Pentland A (eds), *Computer Vision in Human–Machine Interaction*. Cambridge University Press, pp. 191–216.

Chastain T, Kalay YE and Peri C (2002) Square peg in a round hole or horseless carriage? Reflections on the use of computing in architecture. *Automation in Construction* **11**, 237–248.

Cicognani A and Maher ML (1997) Design Speech Acts. "How to Do Things with Words" in Virtual Communities. In *CAAD futures 1997*. pp. 707–717.

Clay SR and Wilhelms J (1996) Put: language-based interactive manipulation of objects. *IEEE Computer Graphics and Applications* **16**, 31–39.

Coroado L, Pedro T, D'Alpuim J, Eloy S and Dias MS (2015) VIARMODES: visualization and interaction in immersive virtual reality for the architectural design process. In Martens B, Wurzer G, Grasl T, Lorenz W and Schaffranek R (eds), *Real Time – Proceedings of the 33rd eCAADe Conference – Volume 1*. Vienna, Austria, pp. 125–134.

Crawley EF, Malmqvist J, Östlund S, Brodeur DR and Edström K (2014) *Rethinking Engineering Education: The CDIO Approach*, 2nd Edn. Basel: Springer International Publishing.

Devlin K (1990) An examination of architectural interpretation: architects versus non-architects. *Journal of Architectural and Planning Research* **7**, 235–244.

Dong A (2005) The latent semantic approach to studying design team communication. *Design Studies* **26**, 445–461.

Eckert C and Stacey M (2000) Sources of inspiration: a language of design. *Design Studies* **21**, 523–538.

Eckert C, Kelly I and Stacey M (1999) Interactive generative systems for conceptual design: an empirical perspective. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **13**, 303–320.

Fish J (2004) Cognitive catalysis: sketches for a time-lagged brain. In Goldschmidt G and Porter W (eds), *Design Representation*. pp. 151–184. London: Springer-Verlag London Limited.

Furnas GW, Gomez LM, Landauer TK and Dumais ST (1982) Statistical semantics: How can a computer use what people name things to guess what things people mean when they name things? In *Proceedings of the 1982 conference on Human factors in computing systems – CHI '82* (pp. 251–253). New York, New York, USA: ACM Press.

Gabriel GC and Maher ML (2002) Coding and modelling communication in architectural collaborative design. *Automation in Construction* **11**, 199–211.

Ghomeshi M and Jusan MM (2013) Investigating different aesthetic preferences between architects and non-architects in residential façade designs. *Indoor and Built Environment* **22**, 952–964.

Gifford R, Hine DW, Muller-clemm W and Shaw KT (2002) Why architects and laypersons judge buildings differently: cognitive properties and physical bases. *Journal of Architectural and Planning Research* **19**, 132–148.

Glock F (2009) Aspects of language use in design conversation. *CoDesign* **5**, 5–19.

Goel V (1995) *Sketches of Thought*. Cambridge, Massachusetts: MIT Press.

Goldschmidt G (2004) Design representation: private process, public image. In *Design Representation*. London: Springer London, pp. 203–217.

Groat L (1982) Meaning in post-modern architecture: an examination using the multiple sorting task. *Journal of Environmental Psychology* **2**, 3–22.

Gross MD (1996) The electronic cocktail napkin – a computational environment for working with design diagrams. *Design Studies* **17**, 53.

Harrison S and Minneman S (1996) A bike in hand: a study of 3-D objects in design. In Cross N, Christiaans H and Dorst K (eds), *Analysing Design Activity*. John Wiley & Sons, pp. 417–436.

Herbert DM (1988) Study drawings in architectural design: their properties as a graphic medium. *Journal of Architectural Education* **41**, 26–38.

Herold J and Stahovich TF (2011) Using speech to identify gesture pen strokes in collaborative, multimodal device descriptions. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **25**, 237–254.

Jonson B (2005) Design ideation: the conceptual sketch in the digital age. *Design Studies* **26**, 613–624.

Jørgensen U (2007) Historical accounts of engineering education. In *Rethinking Engineering Education: The CDIO Approach*. Boston, MA: Springer US, pp. 216–240.

Khan S and Tuncer B (n.d.) 3D CAD modeling using gestures and speech: Investigating CAD legacy and non-legacy procedures. (*Forthcoming*).

Khan S and Tuncer B (2017) Intuitive and effective gestures for conceptual architectural design. In *ACADIA 2017: Disciplines & Disruption [Proceedings of the 37th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)* (pp. 318–323). Cambridge, Massachusetts.

Khan S, Rajapakse H, Zhang H, Nanayakkara S, Tuncer B and Blessing L (2017) GesCAD. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction – OZCHI '17* (pp. 402–406). New York, New York, USA: ACM Press.

Köpsel A and Bubalo N (2015) Benefiting from legacy bias. *Interactions* **22**, 44–47.

Lawson B and Loke SM (1997) Computers, words and pictures. *Design Studies* **18**, 171–183.

Lenau T and Boelskifte P (2005) Verbal communication of semantic content in products. In Binder T, Grogh PG, Redström J and Mazé R (eds), *Nordes*

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press

Conference "In the Making". Copenhagen: Royal Danish Academy of Fine Arts, School of Architecture, pp. 11–23.

Llinares C and Iñarra S (2014) Human factors in computer simulations of urban environment. Differences between architects and non-architects' assessments. *Displays* **35**, 126–140.

Llinares Millán MDC, Iñarra S and Guixeres J (2018) Design attributes influencing the success of urban 3D visualizations: differences in assessments according to training and intention. *Journal of Urban Technology* **25**, 39–57.

Logan GD and Radcliffe DF (2004) Impromptu prototyping and artefacting: representing design ideas through things at hand, actions, and talk. In Goldschmidt G and Porter WL (eds), *Design Representation*. pp. 127–148. London: Springer-Verlag London Limited.

Luck R (2013) Articulating (mis)understanding across design discipline interfaces at a design team meeting. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **27**, 155–166.

Malizia A and Bellucci A (2012) The artificiality of natural user interfaces. *Communications of the ACM* **55**, 36.

McNeill D (1992) *Hand and Mind: What Gestures Reveal About Thought*. Chicago: Univ. of Chicago Press.

Menegotto JL (2015) Computer-Aided architectural design futures. *The Next City – New Technologies and the Future of the Built Environment* **527**, 329–347.

Mignonneau L and Sommerer C (2005) Designing emotional, metaphoric, natural and intuitive interfaces for interactive art, edutainment and mobile communications. *Computers and Graphics* **29**, 837–851.

Minneman SL and Harrison SR (1998) Negotiating right along – An extended case study of the social activity of engineering design. In Duffy AHB (ed.) *The Design Productivity Debate*. Berlin, Germany: Springer-Verlag, pp. 32–50.

Morris MR (2012) Web on the wall: insights from a multimodal interaction elicitation study. In Orit Shaer, Chia Shen, Meredith Ringel Morris and Michael Horn (eds), *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, 95–104.

Morris MR, Danielescu A, Drucker S, Fisher D, Lee B, Schraefel M and Wobbrock JO (2014) Reducing legacy bias in gesture elicitation studies. *Interactions* **21**, 40–45. doi: 10.1017/CBO9781107415324.004

Najari A, Dubois S, Barth M and Sonntag M (2016) From Altshuller to Alexander: towards a bridge between architects and engineers. *Procedia CIRP* **39**, 119–124.

Nanjundaswamy VG, Kulkarni A, Chen Z, Jaiswal PSSS, Verma A and Rai R (2013) Intuitive 3D Computer-Aided Design (CAD) System With Multimodal Interfaces. In *ASME 2013 International Design Engineering Technical Conferences (IDETC) and Computers and Information in Engineering Conference (CIE)*. Portland, Oregon, USA: ASME.

Neal JG, Thielman CY, Dobes Z, Haller SM and Shapiro SC (1989) Natural language with integrated deictic and graphic gestures. In *Proceedings of the Workshop on Speech and Natural Language – HLT '89*. Cape Cod, Massachusetts: Association for Computational Linguistics, p. 410.

Oak A (2011) What can talk tell us about design?: analyzing conversation to understand practice. *Design Studies* **32**, 211–234.

Oh J, Stuerzlinger W and Danahy J (2006) SESAME: towards better 3D conceptual design systems. In Carroll JM, Bødker S and Coughlin J (eds), *Proceedings of the 6th Conference on Designing Interactive Systems, DIS '06*. New York, USA: ACM, pp. 80–89.

Oviatt S (1999) Ten myths of multimodal interaction. *Communications of the ACM* **42**, 74–81.

Oviatt S, DeAngeli A and Kuhn K (1997) Integration and synchronization of input modes during multimodal human–computer interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems – CHI* **97**, 415–422.

Podehl G (2002) Terms and measures for styling properties. In *DS 30: Proceedings of DESIGN 2002, the 7th International Design Conference*. Dubrovnik, Croatia: Sveacilisna Tiskara, pp. 879–886.

Poggenpohl S, Chayutsahakij P and Jeamsinkul C (2004) Language definition and its role in developing a design discourse. *Design Studies* **25**, 579–605.

Purcell T (1996) The data in design protocols. The issue of data coding, data analysis the development of models of the design process. In Cross N, Christiaans H & Dorst K (eds), *Analysing Design Activity*. New York: John Wiley & Sons, pp. 225–252.

Quek F, McNeill D, Bryll R, Duncan S, Ma X-F, Kirbas C, McCullough KE and Ansari R (2002) Multimodal human discourse: gesture and speech. *ACM Transactions on Computer-Human Interaction* **9**, 171–193.

Salisbury MW, Hendrickson JH, Lammers TL, Fu C and Moody SA (1990) Talk and draw: bundling speech and graphics. *Computer* **23**, 59–65.

Segers N and Leclercq P (2007) Computational linguistics for design, maintenance, and manufacturing. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **21**, 99–101.

Stacey M and Eckert C (2003) Against ambiguity. *Computer Supported Cooperative Work (CSCW)* **12**, 153–183.

Tan YY (2014) English as a "mother tongue" in Singapore. *World Englishes*, **33**, 319–339.

Tunçer B and Khan S (2018) User defined conceptual modeling gestures. In Lee J-H (ed.) *Computational Studies on Cultural Variation and Heredity*. Singapore: Springer, pp. 115–125.

Varshney S (1998) Castle in the Air: a strategy to model shapes in a computer. *Proceedings of the Asia Pacific Conference on Computer Human Interaction (APCHI 98)* **3rd**, 350–355.

Walther J, Robertson BF and Radcliffe DF (2007) Avoiding the potential negative influence of CAD tools on the formation of students' creativity. In Søndergaard H and Hadgraft R (eds), *Proceedings of the 18th Conference of the Australasian Association for Engineering Education (AaeE)*. Melbourne, Australia: University of Melbourne, pp. 1–6.

Wang R, Paris S and Popović J (2011) 6D hands: markerless hand-tracking for computer aided design. In Pierce J, Agrawala M and Klemmer S (eds), *UIST 11 Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*. Santa Barbara, CA, USA: ACM, pp. 549–557.

Weimer D and Ganapathy SK (1989) A synthetic visual environment with hand gesturing and voice input. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Wings for the Mind – CHI '89*, **20**. New York, New York, USA: ACM Press, pp. 235–240.

Wiegers T, Langeveld L and Vergeest J (2011) Shape language: how people describe shapes and shape operations. *Design Studies* **32**, 333–347.

Wobbrock JO, Morris MR and Wilson AD (2009) User-defined gestures for surface computing. *Proceedings of the 27th International Conference on Human Factors in Computing Systems – CHI* **09**, 1083.

Zheng JM, Chan KW and Gibson I (2001) Desktop virtual reality interface for computer aided conceptual design using geometric techniques. *Journal of Engineering Design* **12**, 309.

Zhong K, Kang J, Qin S and Wang H (2011) Rapid 3D conceptual design based on hand gesture. In *3rd International Conference on Advanced Computer Control*. pp. 192–197.

Sumbul Khan (née Ahmad) is a Research Scientist at the SUTD-MIT International Design Centre based at the Singapore University of Technology and Design. Sumbul holds a PhD in Architecture from the University of Strathclyde, UK, specializing in computational design. Her credentials include a professional B.Arch. degree from the GGS Indraprastha University, New Delhi, India, and an MSc in Architectural Computing Studies from the University of Strathclyde, UK. Her research interests include generative design, computer-aided design, and human–computer interaction.

Bige Tunçer is an associate professor and the associate head of pillar at the Architecture and Sustainable Design Pillar of the Singapore University of Technology and Design (SUTD). At SUTD, she founded the Informed Design Research lab, which focuses on data collection, information and knowledge modeling, and visualization, for informed architectural and urban design. She leads and participates in various research projects in evidence informed design. Her research has been widely published internationally in books, journals, and conference proceedings. She has taught many design computation and design studio courses to undergraduate and graduate students.

https://doi.org/10.1017/S0890060419000015 Published online by Cambridge University Press