A multimodal study of augmented reality in the architectural design studio

Alejandro Veliz Reyes[1]

Abstract

Architectural design representations comprise physical (models, drawings) and digital media (3D models). Multimodal combinations of physical and digital representations are commonplace in professional contexts such as design reviews or building sites, and 'making sense' of these assemblies requires complex socio-cognitive processes and professional expertise. In that context, this paper follows a 'sense-making' approach, and theorises the impact of multimodal representations in architectural design tutorials. Hybrid physical/digital representations have been built using Augmented Reality (AR) and utilised in dialogues between students' groups and design tutors. Design tutorials have been documented through video, workshops, observational notes and interviews to generate a grounded theory through iterative and reflexive coding. Following a multimodal approach, the paper frames AR as an enabler and mediator of design communication, evidenced through multimodal choreographies of physical/digital media, speech and the individual and collective performances of participants' bodies and actions in space. The resulting theory is composed of 7 concepts outlining the impact of AR in multimodal architectural communication, including the major constructs 'AR-mediated Interaction' and 'Augmented Pedagogies'. The paper outlines this conceptual taxonomy and provides fieldwork evidence supporting a methodological shift from technology-focused to sense-making observations of technology in design activity.

Keywords Design communication · Augmented reality · Multimodality · Digital design

Introduction

In architectural design research, the terms 'multimodality' or 'multimodal' often refer to collections of 'types' of design information and knowledge representation such as images or text related to visualisation or information management systems (Aksamija & Iordanova,

✉ Alejandro Veliz Reyes
    alejandro.velizreyes@plymouth.ac.uk

[1] School of Art, Design and Architecture, University of Plymouth, Plymouth, United Kingdom

2010). Literature also refers to 'modes of designing', and although not explicitly defined, the term 'mode' has been utilised to provide descriptive frameworks of design tasks and the different ways those tasks are organised: 'drawing, writing, thinking, listening and examining' (Akin & Lin, 1995) or 'gathering information, sketching, and reflecting' (Cross, 2001). This task-based approach has been foundational to contemporary design research and has substantially contributed to the study of the complexities of creativity and its physical embodiment (Gero & McNeill, 1998; Purcell & Gero, 1998; Suwa et al., 1999). Architectural design is, however, a complex and distributed socio-cognitive process, involving environments of practice and learning such as the design studio (Schön, 1985; Webster, 2008), professional cultures and traditions (Cuff, 1991), historical forms of architectural representation (Hewitt, 2014), institutionalised power dynamics and 'rites of passage' embedded in both learning and professional frameworks (Dutton, 1987) and intricate dynamics balancing social and technological aspects of design activity (Picon, 2022). Isolating design practice from its contextual complexities can present a partial account of how design unfolds, providing limited views and interpretations of design tasks and their associated modes of communication (e.g., drawing, sketching) leaving aside relevant emotional, social or cultural aspects (Webster, 2008) of design activity. This paper 'complexifies' (Tsoukas, 2017) design activity in the context of design tutorials acknowledging these broader aspects, and theorises the broader implications of multimodal representations in architectural design tutorials.

Nomenclature around design communication is contested in the literature, although this paper focuses specifically on design tutorials (either individual or in groups). These are considered a form of learning distinctively different from the more formal design reviews or design juries (Oh et al., 2013). Design tutorials are typical for both learning and professional environments (often called the design feedback) and comprises formative feedback, design iteration, reflection, and discussions on further work (Tahsiri, 2020) and a combination of communication modalities such as gesture, drawing and speech (Oh et al., 2013). Additional evidence indicates how design tutorials enact an acculturation process, facilitating learners' transitions towards professional behaviour and skills (Taneri & Dogan, 2021; Thompson, 2019). Instead of a task-focused observational framework, however, design tutorials are here documented, transcribed and analysed from a multimodal perspective. Multimodality is a research approach stemming from social semiotics and communication, mainly through the work of theoreticians such as Kress, van Leeuwen and Jewitt (Jewitt, 2013, 2014; Kress, 2010; Kress & Leeuwen, 2001). The goal of multimodal research is to observe the relational ties across different modes of communication aiming to reveal meaning-making processes in socially and culturally shaped environments: instead of a task-orientated approach, multimodality defines a mode of interaction as a 'socially and culturally shaped resource for making meaning' (Jewitt, 2013). It is, for instance, in the professional culture of the architectural design studio where a floor plan comprises meaning (for example, spatial planning intents) shaped by a collection of semiotic resources: linewidth, colour or hatching patterns, all understandable and communicable in the context of architectural professional practice (Fig. 1).

Building upon task-based analyses, a meaning-making approach is a necessary response to the changing nature of contemporary communication (Kress & Leeuwen, 2001) as it relies on the evolving nature of texts (and text-making) and language: new digital practices and technologies are radically changing the way we communicate and represent design knowledge. Designers indeed model, talk, gesticulate, and draw to a high level of complex-

Fig. 1 An architectural design review including assemblies of physical (print, models) and digital (drawings, building visualisations) media, as well as interaction between participants through speech and gestural communication

ity, and design practice comprises observing, developing and implementing modes of design interaction: the use of screens as design interfaces, the use of computers to create design representations and the relevance of visual means of communication would not be odd to any architectural designer today. For example, recent research has documented emerging socio-cognitive frameworks from interactions with novel technologies such as architectural robotics (Smithwick & Sass, 2014). In terms of design interaction, studies such as Mewburn's (2009, 2012) describe how specific modes of communication such as gestures are enacted in 'an architectural way', with embodied meanings highly reliant on specific contexts of architectural studio culture in which those gestures operate. Industry-wide, recent digitally-enabled discourses such as paperless architecture, digital fabrication and manufacturing, or building information modelling heavily rely on complex assemblies of tools and methods of design and data representation. Previous research on the multimodal nature of architectural design communication has focused on the observations of specific modes of interaction such as gestures (Mewburn, 2009, 2012), attributes of a design development process such as form (Luck, 2014), or specific activities rooted in design studio culture, such as the design jury (Murphy et al., 2012). Although not analysed from a multimodal perspective, related recent projects such as (Dorta et al., 2016l, 2018; Sheldon et al., 2019; Smithwick & Sass, 2014) do provide evidence of the benefits and impacts of digitally-mediated social, gestural and vocal communication in design.

Within this research framework, the Methodology section of this paper describes the implementation of a multimodal study to better understand the role of multimodal design representations in design tutorials, acknowledging not only formal and technical aspects of design communication, but more broadly the emotional, social and collaborative dimensions of design activity. The paper will focus on the use of Augmented Reality (AR) technology as a methodological pathway to investigate multimodal design communication (Veliz Reyes, 2016), due to its capacity to merge and visualise both physical (e.g., drawings, prints,

physical models) and digital (e.g., images, digital drawings) design representations (Alp et al., 2023). According to the virtuality continuum by Milgram and Kishino (1994) AR corresponds to the display of digital information within a real environment, and has been utilised for various applications in architecture and construction (design visualisation, construction management, building simulation, among others) at different stages of the design-to-construction process (Chi et al., 2013). More recent scoping reviews identify design reviews and communication as key areas of research potential for AR/VR technologies (Delgado et al., 2020; Noghabaei et al., 2020; Sidani et al., 2021) within the broader AEC sector, aligned with a stream of recent research scoping and evaluating AR applications in design studio environments (Alp et al., 2023; Qureshi, 2019) and codesign activity (Gül, 2018).

Prior to the final discussion and conclusions, the results section describes a grounded theory that describes the impacts of AR in design tutorials. The impact of these results lies on the multimodal evidence of broader socio-technical complexities in design tutorials. For instance, the theory captures the performative nature of design technologies in design tutorials, shifting and reorganising power and bodily dynamics within the studio environments. Likewise, the AR-mediated pedagogies enable shifting degrees of expertise, differentiations between individual and collective modes of design interaction, and complex multimodal layers of design communication not otherwise available through speech or sketching only. These results are expanded through a collection of 7 core concepts grouped on 2 core categories outlining the resulting theory.

Methodology

To achieve this, a grounded theory approach has been constructed to conceptualise the role of AR in design tutorials using a wealth of information sources including on-site video documentation, written memos, and observational work through sketching, AR modelling workshops and participant observations, and follow-up semi-structured interviews with students and design tutors. This process, including workshops, has taken place throughout a year in 2 Architecture programmes in the United Kingdom prior to data collection, and comprised a series of ethical considerations to document, archive and analyse design interaction within studio environments, testing an array of AR visualisation engines prior to fieldwork visits, and outlining design briefs and modelling strategies for AR deployment in design studios. Such contextual grounding enables the production of a grounded theory directly sourced from the context it describes (Glaser & Strauss, 1967). Grounded theory aims to build a systematic line of inquiry and develop new knowledge 'about a particular social phenomenon' (Goulding, 1999) through the development of a descriptive theory. Grounded theory was originally developed in the context of nursing studies, with a focus on professional activity and complex social and technical environments (Glaser & Strauss, 1967). More recently in architectural research, grounded theory has supported the emergence of new 'slants of knowledge' (Zamani & Babaei, 2021) acknowledging the socio-technical complexities of design practice in areas such as participatory codesign (Burke & Veliz-Reyes, 2021) and design innovation (Moslehian et al., 2022).

A variety of methods have fed the theory construction process, enabling the components of the theory (concepts) to emerge through constant comparative analyses (coding) and participants' themselves contributing to and influencing data interpretation and co-con-

struction, guiding the research process from broader and exploratory instances up to more focused and refined lines of inquiry as the research process proceeds. Codes and categories emerge, then, directly from the data instead of from pre-existing theoretical constructions (Urquhart, 2007) through a coding process based on constant comparative analyses. This approach enables, additionally, for broader socio-cultural contexts of practice, technologies (Urquhart & Fernández, 2013) and academic training to influence the data, enabling the 'social reality' (Thornberg & Charmaz, 2014) of the design studio to mediate the interpretation of data and its theoretical derivation.

AR implementation in design studios

AR technology used in this research is marker-based – it involves the use of printed images (markers) to track the location of 3D models in space using a camera; once the marker is recognised by a camera, a 3D model is displayed on screen and tracking happens in real time. This type of AR visualisation allows designers to blend virtual (e.g., 3D models) and physical (e.g., architectural models and drawings) media in single, multimodal representation devices which can be visualised in real-time by multiple users on any standard computer screen (Fig. 2) using a camera. Other modes of mixed media representation such as Virtual Reality often work based on individual interfaces such as Head Mounted Displays, restricting the potential for group discussion and collective real-time engagement with design representations (Dorta et al., 2016).
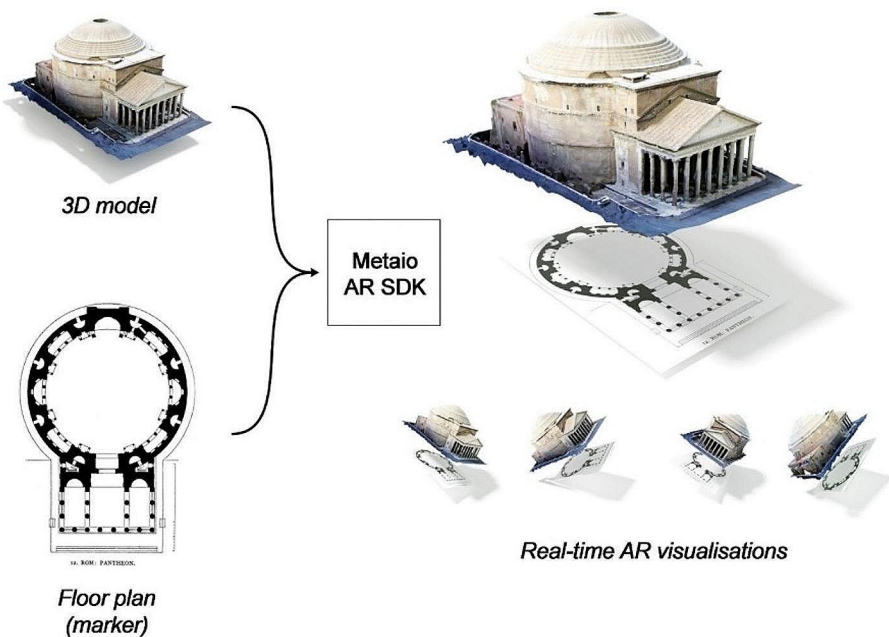
Fig. 2 Sample AR system where a 3D model (indicating volume, materials, proportions) is aligned and simultaneously visualised with the building's floor plan (indicating spatial distribution) acting as a trackable marker

Although architecture students are often skilled on the modelling and visualisation of 3D geometry, the research required co-developing AR visualisations in order to deal with technical challenges and the use of a new visualisation environment (Metaio SDK) for AR modelling. Initial workshops have been conducted at 2 European universities and later on rolled out across a larger cohort of students in two design studios in an Italian school of Architecture with a diverse cohort of international students. Participants utilised AR to communicate their design intents (e.g., volumetric models, images), and developed strategies for design conversations with tutors (e.g., interchangeable markers). The delivery of AR modelling workshops followed not only skills development but additionally discussions regarding its applicability and potentials for design communication and visualisation.

Video documentation of design communication

The video data described in this paper has been recorded over a two-week fieldwork activity in two undergraduate design studios in an Italian school of Architecture. Specifically, this was embedded in the delivery plan of an international cohort of architecture students, involving 2 design studio courses of 22 students' each. Group tutorials ranged between 3 and 5 students each, typically with 2 design studio tutors. Throughout fieldwork, near 32 h of video recording in design studios have been captured, including design interactions with and without the use of AR, and interviews with participants. After 2 AR training workshops, a full day of design tutorials (7 h) has been initially recorded without the use of AR, in order to contextualise and better understand the students' and tutors' dynamics within the studio environment. This was followed by 2 teaching days (near 10 h) of recording design tutorials with and without the use of AR. During design tutorial sessions, students and tutors transitioned through a range of design representations, both digital and physical, as well as AR representations. These sessions all contributed to the coding process, and specific actions for micro-analysis were additionally selected and transcribed over 8 min of video recording, including gestural actions under 1 s of duration. Last, a range of interviews has been recorded with studio participants in order to capture additional insights, confirm observations, record their feedback on the use of AR, and more substantially contextualising (Jewitt, 2012) studio activity.

During video recording, participants continuously shift across a range of media for design representation and communication, and a detailed transcription and microanalysis of multimodal communication has been conducted on dialogues involving the use of AR in design tutorials through the annotation of speech, gestural and body movements' modes of communication (Fig. 3). Such nuanced and non-verbal communication allow for a rich data source including the nature of communication patterns and micro-gestural cues not otherwise visible through audio or observational work only. In specific situations drawn out from the data, video frames have been sketched and redrawn to a clearer and larger resolution in order to outline communication patterns with more detail, and speech has been transcribed by adapting a Jefferson notation model to capture details such as pauses and overlaps between speakers (Atkinson & Heritage, 1985).
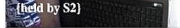
Group 2 – AR-based dialogue 1 (00:03:34)

	Dialogue	Gestures	{Screenshot}
	--- Beginning of the conversation ---	This dialogue is the beginning of a design critique for the group 2. Is focused on the setup of the AR models on the screen and how the technology operates.	
I1	Yes yes yes yes (.) wait wait		
S1	(-) is going well, a (red model) [I1: ok, ok] (-)		
S2	(let's see)		
S1	move it a bit	Both student and instructors are {watching at the AR model} in a laptop screen {held by S2}.	
I1	excuse me		
S1	yes (9)	I1 takes the marker and starts moving it, {overlapping} the 3D model with the {printed floor plan}. S2 tries to {modify the AR scene} in order to improve the visualisation of the 3D model. During the pause (9) students mix several spare words (not audible due to overlaps) to make fun of the situation of holding the laptop to visualize the model properly.	
I2	you know, that you can eventually use a webcam?		
S1	eh		
I2	so you don't need to	"eh" in approval tone.	
S2	yes		
I1	you have to (-) so the model (fits) (.) good (10)		

Fig. 3 Sample data transcript of video recordings including speech, gestural communication and isolated video frames

Results

The development of theory in IT research can take multiple forms: concepts, models, instances and methods (March & Smith, 1995). Through constant comparative analysis and coding, broader observations are refined, tested, and fed by additional patterns and sources, and constructed into emerging theoretical concepts until a point of saturation – that is, when additional data fails to further advance the definition and scope of such concepts (Charmaz, 2006). The theory presented below outlines a conceptual understanding of AR-mediated communication in design studios and develops two core major constructs addressing an emerging characterisation of 'AR-mediated Interaction' in design studios, as well as the socio-cultural, professional and collective dynamics resulting from those interactions in a design communication environment (i.e. 'Augmented Pedagogies') within a multimodal discourse. These major constructs cluster a series of emerging theoretical concepts, outlined below and illustrated through extracts of video documentation and the voice of the research participants.

AR-mediated interaction

In constructing an account of AR-mediated interactions, participants both reflect on the potential and creative opportunities for AR in design studios, and project those opportunities on their own design conversations using AR models reflective of their design goals and digital skills. The mediation of AR in design communication is, then, not only 'representational' in nature, but prompts new conceptual and projective functionalities (Klaasen, 2002) through the visualisation of 3D models and abstract information such as distances and

layout tests (Fig. 4). Through these functions, designers were able to not only communicate intent but to establish additional structural features (Veliz Reyes et al., 2012) mediating dialogue such as measuring distances on-screen, combining image, volume and print media, or exploring different levels of detail in AR visualisations. As a result, conceptualising AR-mediated interaction indicates an acknowledgement of a socio-technical system and the conditions upon which that system is meaningful to the design discussion, rather than a purely technological sense of software or solution deployment.

These nuances, however, are largely facilitated by the ability of designers to move across different aspects of a design project, with participants engaging in dialogue intermittently and a variety of roles emerging throughout the discussion – including lone designers, quietly in the periphery of the core discussion yet somehow engaging with design representations and AR visualisations. Within a detailed coding of those micro-stories in the datasets, the graduation between solo (interaction of designers with technology) and collective dynamics (interaction of designers mediated by technology) has been established as a conceptualisation of this core construct.

Solo interaction

On one end of this gradient is the conceptualisation of solo interactions between a designer and AR-models. Although this concept is, arguably, peripheral to the collective nature of a group discussion, solo interactions reveal fringe behaviour not otherwise able to be documented through speech recording, including role-playing dynamics (i.e. the group leader, participants isolating themselves from the discussion) and opportunities for novice designers to intermittently engage with an array of topics and design issues as the conversation moves on. The extent to which users engage with individual practices changes across groups and participants, yet is present in all design dialogues. Either driven by curiosity, novelty or reflective behaviour, the extent of solo interaction with technology speaks to a substantial peripheral activity underscoring the dynamics of participation and 'collectiveness' of design tutorials. Moreover, solo interactions are not bounded by levels of expertise (as shown in Fig. 5, with a teaching assistant as the main participant).

Solo interactions and their broader conversational context reveal a contrast between the instructional nature of the design studio – often portrayed in literature by the ability of an expert designer to manage the direction of a tutorial covering a range of issues, and a novice
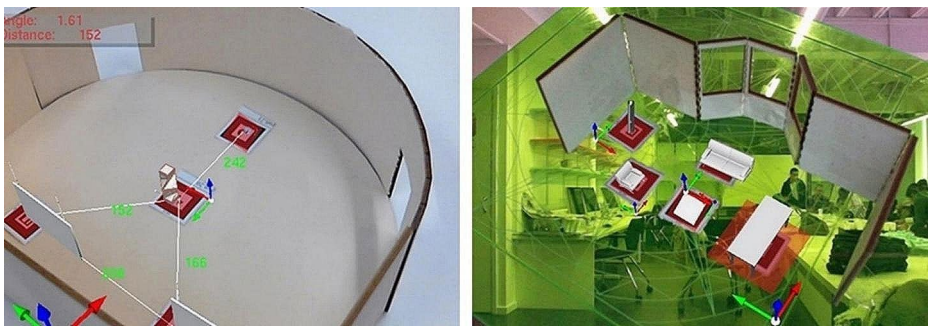
Fig. 4 Layout planning using AR marker tracking and real time distance measurements (screenshots from camera views)

Fig. 5 Sequence of 2 video frames demonstrating an instance of solo interaction. Left: Students (standing up) explain their design to an instructor (left edge). Right: A teaching assistant explores an AR model in isolation as the conversation unfolds around him
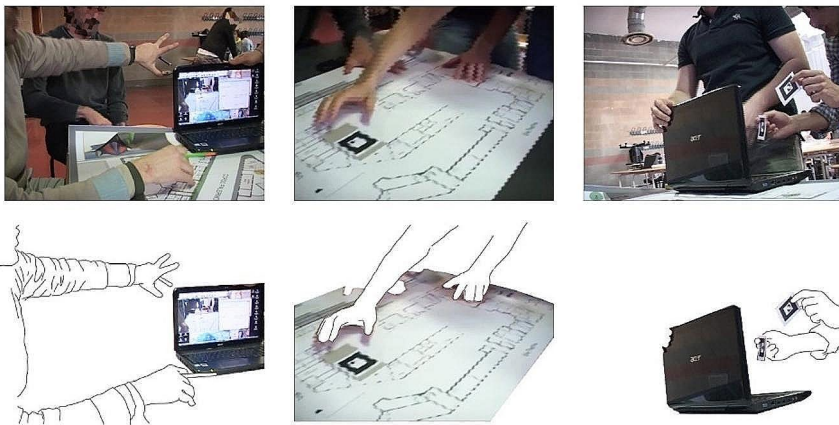
Fig. 6 Range of gestural activity performed with AR in the context of design dialogues including (from left to right) pointing at a digital model on screen, overlapping a marker with a printed floor plan, or combining markers on-screen

designer discussing and reflecting upon those 'design moves' – and more nuanced, isolated behaviours enabled by new representational opportunities. In the case of AR, this technology requires users to not only visualise design information on a screen, but to perform an activity in the physical and social space of the design dialogue (e.g., picking up a marker, moving it in the space) in a 'silent' socio-technical engagement mediated by a combination of physical and digital media (Fig. 6).

Collective interaction

As a counterpart to a conceptualisation of solo interaction, this construct encapsulates more broadly the way participants collectively engage with one another through AR visualisa-

tions. This type of situations is at the core of group discussion and visible through a range of sequentially orchestrated modes of communication including speech, gestural communication and actions with and upon models and representations, and embodies the group's capacity to enact key principles of design studio acculturation, such as somewhat problematic authority and power dynamics through feedback and iterative revision requirements targeting both the students' design process and the design outputs:

'(…) what you showed to me is absolutely very good, but I am personally very worried for the starting scheme. In a way you have a starting scheme and then you have options. Your options are good but the starting scheme is not very good, so it's like to say "producing very good cheese with very bad milk". You get my point?' [tutor participant, translated from Italian].

Throughout the data, a large presence of collective interaction mediated by AR focuses on two key conversational spaces: discussions about design decision-making, and feedback – both spaces closely intertwined yet identifiable in the data through detailed video speech transcription. Multimodal interaction then unfolds through an intricate pattern of design enquiry, discussion and feedback - such complex dynamic, however rich, is elusive in nature as moves across different degrees of participation (from 2 to 5 participants), the performance of bodies in the space (gestures and navigation across multiple models and design representations, and actions such as sketching), and overlapping speech and gestural communication patterns. Here, an awareness of the operation of design studio culture and the participant researcher stance during observational and video documentation work have proven key to identify multimodal interaction and interpret those culturally-rooted dynamics:

The tutor balances a constructive dialogue with students, together with a strong and explicit guidance in terms of design feedback and further design 'moves':

Instructor 1   *yes yes yes yes (.) wait wait*
Student 1 *(-) is going well, a (red model) (-) [Instructor 1: ok ok]*
Student 2 *(let's see)*
Student 1 *move it a bit*
Instructor 1   *excuse me*
Student 1 *yes (9)*

For a reader unfamiliarised with the field not much can be inferred from the dialogue – something 'is going well' and that might be referring to a 'red model'. However, a multimodal interpretation allows to dissect further the meaning of this sequence (Fig. 7) by providing a communication context including gestural, positional, and environmental information. The tutor asks students to 'wait' in order to visualise an AR model on screen (Fig. 7a) while a student holds a laptop (Fig. 7b). After a student points at a model ('is going well') the tutor wishes to check the model on screen ('let's see') (Fig. 7c) and ends up picking the marker ('excuse me') and manipulating the model himself for silent 9 s (Fig. 7d).

This sequence of actions takes place over 27 s and a series of relevant actions can be outlined, including the manipulation of an AR representation, the definition of roles and logistic conditions (i.e. 'holding the laptop'), or the construction of an AR representation to convey an idea of volumetric intent. The distributed nature of these interactions additionally highlights roles and responsibilities emerging from design dialogue, both individually (e.g., the group leader, the lead design tutor) and collectively (e.g., groups taking responsibility
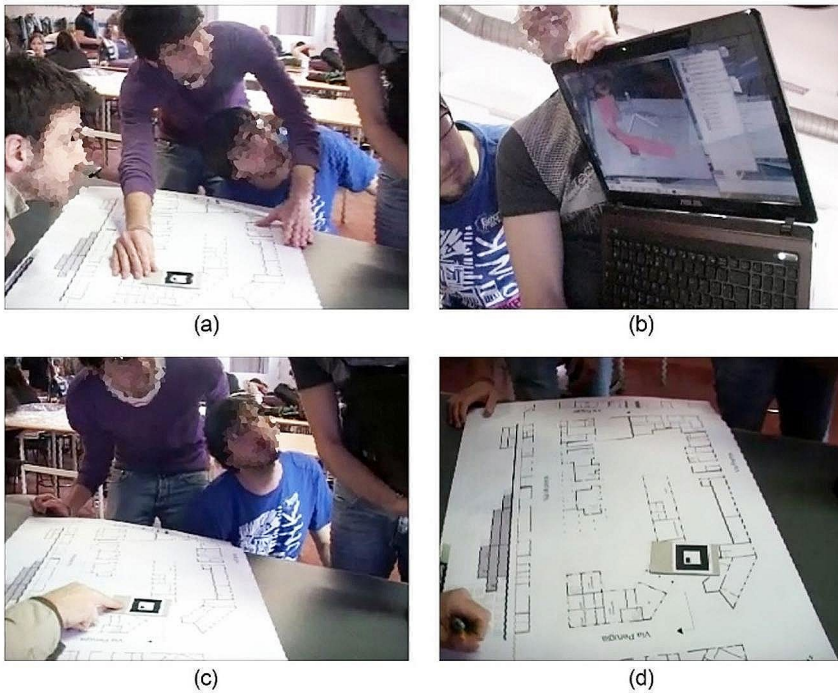
Fig. 7 Video frames relevant to describe the interaction between a group of students and an instructor, sequenced indicated in the text (a, b, c, d)

for design decisions). Collective interaction acts, then, within this 'orchestration of incidents' rather than through isolated situations, and yet the leading role of the tutor remains as a driver of the communication pattern – in this case, with limited experience using and visualising AR representations, underscoring the tutors' ability to shift and manage the discussion while acknowledging the novelty or AR representations.

Augmented pedagogies

Although issues related to skills development and troubleshooting are prevalent in the use of digital design tools in studio settings (Veliz Reyes, 2016), the complex communication observed in the video records suggest a more intertwined, critical and profound layer of analysis. Architectural education comprises what Dutton calls a 'signature pedagogy' (Dutton, 1987) which does not refer only to didactic strategies and curriculum but replicates and reinforces ideologies of practice through socio-cultural and power dynamics. Participants of this research contributed to co-construct the ways in which these dynamics are mediated and enabled by digital technology. While previously detailed concepts (i.e. solo and collective interactions) focus on the role of AR in design dialogue, this section attempts to explain how AR visibilises the studio signature pedagogies and broader emotional and cognitive patterns of activity. Concepts within this major construct include the definition of 'organisational shifts', 'emotional engagement', 'cognitive engagement', 'troubleshooting' and 'technology affordances'.

Organisational shifts

In design studio literature, the authority and standing of more experienced designers is not only uncontested but portrayed as a tenet of the professional and learning ethos of the design studio community, and a gatekeeping mechanism for more novice designers to 'become architects'. Such de-facto assumption of competence is a foundational aspect of the theory of reflective practice and is outlined in detail by Schön (1983) on his account of the conversation between a design tutor (Quist) and a student (Petra). In the context of this research and through video records of collective dialogue in groups between 4 and 7 participants, the incorporation of a new technology into their design practice shifts the organisational dynamics of group work by enabling more complex patterns of communication than 1-to-1 instructional conversation, including shifting role-playing patterns (e.g., students speaking from a position of 'the AR modeller', or members of the group dealing with logistics e.g., 'the camera holder') formed on the basis of new representational capabilities within the group (i.e. AR visualisation) and negotiations between well-known (A1 size posters with drawings) and novel AR-based digital representations (Fig. 8).

These shifts are reflected on design dialogues through the enactment of such roles in gestural and speech communication, as well as the re-organisation of bodies in the physical space of the design tutorial, rearranging participants as students respond to different conversational scenarios. These fluid spatial organisations often comprise a shift on the pedagogical and power structures of the tutorial, sometimes leaving the teaching team in a position of passive observers while students navigate and visualise AR representations on-screen. However short, these moments of shifting hierarchy suggest a degree of opportunistic behaviour enabling students to assume a more dominant role from a position of technological competence and navigation of the AR visualisations in front of tutors, allowing them the opportunity to assert some degree of (technological) dominance. This spatialised and nuanced power dynamic contests the broadly accepted position of expertise by design tutors and enables students to explore a, however narrow, space of self-development free of tutors' disciplinary control (Webster, 2008).

Emotional engagement

Previous research suggests a limited understanding of how affective dimensions of design studio relate to broader disciplinary power and acculturation processes (Webster, 2008) and, more broadly, how affective learning activities mediate the progression of a learning process
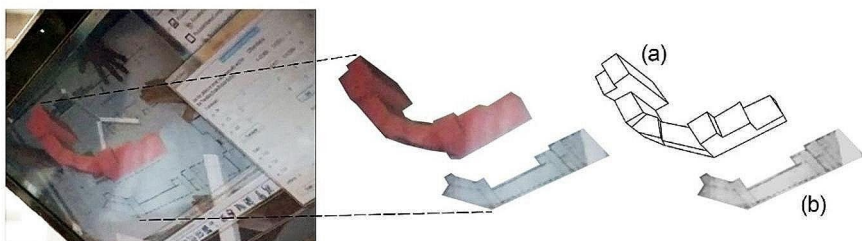

Fig. 8 A video frame of a student's computer screen during a design tutorial showing a (a) digital volumetric model overlapped with (b) a building proposition floor plan

within broader cognitive and disciplinary contexts of digital practice (Picard et al., 2004). Likely a result of IT research focused on systems development, emotional and affective aspects of learning and its relationship with creativity are often minimised, and are here constructed by participants in close intertwines with situations and slices of data referred to demonstrations of motivation and interest through speech ('Instructor: well done adding this cardboard slip') or through multimodal combinations of speech, gestural communication and actions with and upon models (Fig. 9).

Moreover, motivational speech seems to not only reflect predisposition to the use of AR (for instance, how students seemed to be creatively attached to some AR visualisations), but 'sets the tone' for design dialogues and largely mediates a broader array of situations such

Instructor: ah another marker
[Student 1: yes] can we use this?
Instructor: well done adding this cardboard slip

(a)

(b)

(c)
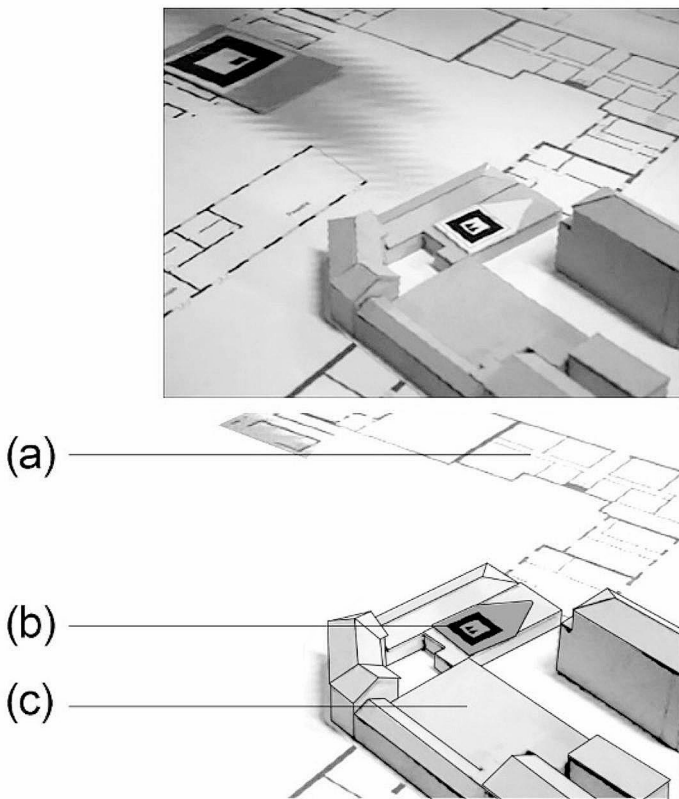
Fig. 9 A students' device to augment a physical model to visualise a range of design options. A base drawing (a) and an interchangeable set of AR markers (b) are used to increase the representational capabilities of a physical cardboard model (c)

as lines of questioning related to key discussion points such as students' spatial and aesthetic choices - a 'microtechnology of power' (Foucault, 1977) utilised to, at best, guiding students throughout the design discussion offering constructive feedback as they 'become architects' and, at worse, problematically coercing students into questioning their spatial or aesthetic design decisions represented through AR visualisations:

> '(…) I really don't understand why you have these disordered, apparently randomized, and completely scattered shapes. With this line, why this line is here? Why do you have this corner? Why do you have this church like this, and this ridiculous … this absolutely ridiculous triangle just because the church is straight and you want to have this circle. This is ridiculous' [tutor during a design conversation, translated from Italian].

Cognitive engagement

Closely intertwined, creativity, learning and cognition have been the foci of substantial educational research including the tendency of conflating 'design thinking' with 'design learning'. Similarly to the concept of 'emotional engagement', aspects of professional and creative cultures of practice are often marginalised from cognitive science (Hutchins, 1995) in favour of neuropsychological or cognate data-driven fields. During fieldwork, the observation of 'how participants think about things' has been facilitated through a multimodal video analysis by framing cognition as a distributed process comprising people and technologies within a socio-cultural context (Hutchins & Klausen, 1996), and constructing a concept of 'cognitive engagement' by identifying patterns of communication indicating how participants ideate, reflect upon or make decisions about design – sometimes under the threshold of consciousness and yet, demonstrating ways to speak, gesticulate and 'think like an architect' (Schön, 1983). In the data, these instances are usually defined by signifiers of thought, reflection and ideation delivered through micro-bursts of multimodal activity such as beat gestures and actions upon models identifiable in a detailed frame-by-frame analysis (Fig. 10) which are typically left aside from broader analyses of design studio pedagogy (Mewburn, 2012) focused on modes of communication such as speech or sketching.

The multimodal performance of participants is, additionally, mediated by AR media and how it allows them to visualise transitions between ideation and representation using both virtual (i.e. digital models) and physical models (i.e. cardboard models, print drawings). Participants display a tendency to associate virtual models to more abstract, speculative, early-stage design information, and to associate physical representations (either drawings or models) with more 'concrete' representations of a more resolved and better-refined design proposal. Such predisposition is not unique to this fieldwork experience and evidence suggests (e.g., professional accreditation criteria) a professionally rooted view of tangible representations as a more valid mode of communication. In this context, AR plays a unique role enabling the simultaneous visualisation of both digital models more closely related to early-stage tasks such as ideation (more speculative, rough volumetric digital sketches) and physical representation of actual built environments with a specific materiality, scale and spatial intent (more concrete and 'building-like') (Fig. 11). This emerging spectrum of representational capabilities is reminiscent of the 'reproduction fidelity' dimension of Milgram

Instructor:
well done adding this cardboard slip (...)

Student:
then we can substitute it   *(like this) (-) (5)*
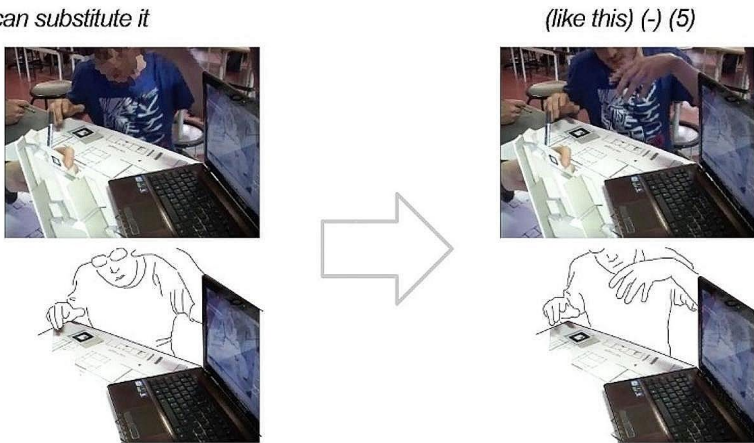
Fig. 10 Gestural action by a student signifying the capacity to substitute a digital model with a beat gesture from his head 'into' a print floor plan

and Kishino's Virtuality Continuum (1994) which, however focused on visualisation hardware, establishes a taxonomy on the basis of user engagement with realism and immersion in mixed reality displays.

Troubleshooting

The incorporation of new digital tools in architectural design contexts is usually framed within an optimisation discourse involving quicker and more effective modes of production to 'better perform' the design/build process (Andia, 2002), either through individual tasks (such as drawing, modelling) or organisational ones (such as managing data and documentation).

Research participants construct this process through a range of temporal frameworks. Within a more immediate time frame, initial steps of AR incorporation in design studios comprise a steep learning curve, experimentation and sorting a range of technical issues (e.g., installation, skills development, hardware requirements). This process is often framed within a 'transitional' narrative enacting the progression from analogue to digital modes of communication. In a longer-term frame, however, participants have defined a more critical appraisal of the challenges to learn and adopt AR technologies in their design work. This process of adoption is both reflected in design dialogue (Fig. 12) and constructed in interviews as a more complex process of trial and error, balancing design skills against design goals and expectations, time management (or lack thereof) and digital skills within architectural education. Troubles using AR technology often reveal these pressures faced by participants: limited time to deliver new skills and content within stringent curricular plans, staffing resource to support the delivery of new skills and digital design technologies,

Instructor:
… an intelligent thing is, for
instance (.) I have a project's floor
plan and a model of the full volume
(-)(.) while I design in the floor plan,
while I am dealing with it, I have the   **other magic model**   *as part of the design process*
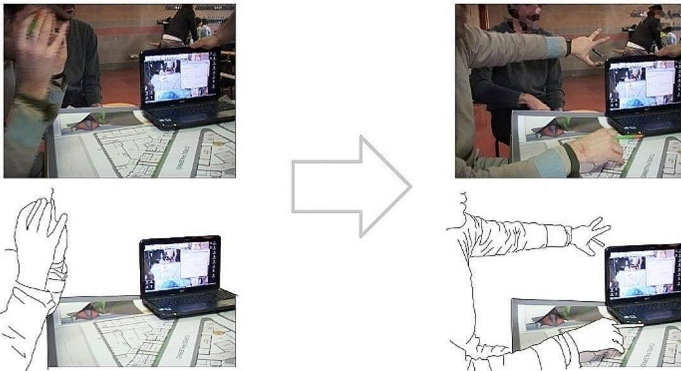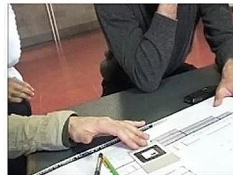

Fig. 11 Gestural action by a tutor (0.27 s) transitioning between print and digital media

Speech (Instructor 1)   Gesture (video frame)

(...) yes but that is a
problem of the model's
visibility, the position,
the georeferencing of
the model. **Also you
didn't place it in the
correct point**, even if it
is difficult I cannot see it
here (...)

*Instructor keeps moving
the marker in the same
plane as the floorplan,
indicating the location
the marker should be at.*

Fig. 12 Multimodal account of a design tutor noting a georeferencing issue with the AR model during a design tutorial

and the willingness of students to learn a new representation tool to report in their professional portfolios. In the context of the fieldwork settings for this research, both Universities are located in countries with strict architecture professional requirements including both an academic degree and a professional qualification compliant with a series of accreditation and legal criteria, often limiting the time and resourcing dedicated to deliver contents

outside canonical architectural design conventions, such as emerging design visualisation technologies.

Technology affordances

Throughout the study, participants did not only develop AR visualisations for specific representational needs, but also speculated with its use and potential outlining opportunities for new modes of visualisations resulting from merging physical and digital media outside the boundaries of the models and representations utilised for design dialogues. Situations framing AR as a vehicle to explore further opportunities instead of a visualisation technology, suggest new design representational practices through the 'availability and affordances' (Kress & Leeuwen, 2001) of AR, stepping beyond the matter at hand in design tutorials (i.e. the discussion and feedback of design work conducted by students) and suggesting 'what is possible to express' instead of what designers are, actually, communicating.

Despite the contested nature of the term 'affordance' in literature and its use to define various relationships between people and technology (Oliver, 2005) within a multimodal discourse, navigating these speculative representational spaces revealed participants' aspirational capabilities and interests, and outlined perceived boundaries of applicability for AR in design dialogues beyond their own skillset.

In this regard, participants utilised the temporal logics of AR to explore these new affordability spaces. Although both physical and digital models are static in nature (no students utilised AR video during fieldwork), the performative nature of AR in design tutorial has a structured 'temporal instantiation' (Jewitt & Henriksen, 2016) through marker recognition, tracking, movement of the marker in the space, its visualisation by participants, and a following discussion. This sequence would be often disrupted through the use of multiple markers, the use of AR images, or malfunctions in the AR tracking during design conversations. These temporal dynamics could not otherwise emerge as a result of static imagery only, and situates AR visualisations within more complex orchestrations of multimodal communication enabling participants to speculate further on the interactive, visual and dynamic nature of AR:

'I haven't seen anything like that; I didn't imagine it exists (.) (-) maybe in the cinematographic dimension, where you can use this virtual creation of elements. It's like a new world, many possibilities of capturing an object that doesn't exist (.) it's exceptional to see and have in my hand an object that is not in my hand' [student participant referring to AR visualisations, translated from Italian].

Discussion and conclusions

This article presents the construction of a theory that describes the impact of AR in design tutorials from a multimodal perspective. The theory is formed by two core constructs which reveal the close and complex links between interaction ('AR-mediated Interaction) and professional culture ('Augmented Pedagogies') in architectural design studios. More profoundly, however, those two major constructs comprise seven concepts dissecting in

detail the micro-communication patterns enacting roles (e.g., 'organisational shifts'), performances (e.g., 'collective interaction') or expectations (e.g., 'technology affordances') of participants in relation to the mediation of AR visualisations during both collective and solo interactions with technology - among other aspects. Relevantly, these concepts assert the influence of technology beyond utilitarian and didactic terms, and more profoundly underscore established power and authority dynamics at play in design studio communication (some of them, established in the literature). The theory encompasses additional aspects of design studio interaction including evidence charting emotional, cognitive or corporeal dimensions (Webster, 2008) of architectural acculturation. Through these conceptualisations the grounded theory provides evidence built around AR utilised "in the wild" and shows how modes of communication are identifiable at micro-scales of analysis of design interaction. This approach aligns with more recent developments across the VR/AR and multimodal research agenda, including areas such as virtual touch (Jewitt et al., 2021; Price et al., 2021) as well as organisational research constructed around the impact of digital technologies within design communities of practice (Stals et al., 2021; Verstegen et al., 2019).

More broadly, a key contribution of this paper is the framing of digital design as a meaning-making socially constructed environment of practice, instead of a task-focused domain of technology development and adoption. By following a multimodal approach, the view of 'analogue versus digital' modes of communication is contested and described as a more nuanced discourse, acknowledging the frictions between the rapidly evolving nature of digital design communication and professionally established institutions and practices. These multimodal 'grammars' are, then, not only descriptions of communication patterns but provide evidence of deeper issues and resistances in design studio culture, and the complexities involved in 'thinking like an architect' through navigating the effects 'with' and 'of' technology (Salomon et al., 1991) in design communication. As a result, multimodal approaches to design communication have the potential to reveal further insights into issues of professional culture, training and technological skills, transitions into the profession, and development of 'digital cultures' within the discipline.

References

Akin, Ö., & Lin, C. (1995). Design protocol data and novel design decisions. *Design Studies*, *16*(2), 211–236. https://doi.org/10.1016/0142-694X(94)00010-B.

Aksamija, A., & Iordanova, I. (2010). Computational environments with multimodal representations of architectural design knowledge. *International Journal of Architectural Computing*, *8*(4), 439–460. https://doi.org/10.1260/1478-0771.8.4.439.

Alp, N. C., Yazici, Y. E., & Oner, D. (2023). Augmented reality experience in an architectural design studio. *Multimedia Tools and Applications*, *82*, 45639–45657. https://doi.org/10.1007/s11042-023-15476-w.

Andia, A. (2002). Reconstructing the effects of computers on practice and education during the past three decades. *Journal of Architectural Education*, *56*(2), 7–13. https://doi.org/10.1162/10464880260472512.

Atkinson, J. M., & Heritage, J. (Eds.). (1985). *Structures of social action: Studies in conversation analysis*. Cambridge University Press.

Burke, R., & Veliz-Reyes, A. (2021). Socio-spatial relationships in design of residential care homes for people living with dementia diagnoses: A grounded theory approach. *Architectural Science Review*, *Special Issue: Architectural Design Science for Dementia*. https://doi.org/10.1080/00038628.2021.1941749.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE.

Chi, H. L., Kang, S. C., & Wang, X. (2013). Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation in Construction*, *33*(August 2013), 116–122. https://doi.org/10.1016/j.autcon.2012.12.017.

Cross, N. (2001). Design cognition: Results from protocol and other empirical studies of design activity. In C. Eastman, M. McCracken, & W. Newsletter (Eds.), *Design Knowing and Learning: Cognition in Design Education* (1 ed., pp. 79–103). Elsevier Science. https://doi.org/10.1016/B978-008043868-9/50005-X.

Cuff, D. (1991). *Architecture: The story of practice*. MIT Press.

Delgado, J. M. D., Oyedele, L., Demian, P., & Beach, T. (2020). A research agenda for augmented and virtual reality in architecture, engineering and construction. *Advanced Engineering Informatics*, *45*(August 2020), 101122. https://doi.org/10.1016/j.aei.2020.101122.

Dorta, T., Kinayoglu, G., & Hoffmann, M. (2016). Hyve-3D and the 3D cursor: Architectural co-design with freedom in virtual reality. *International Journal of Architectural Computing*, *14*(2), 87–102. https://doi.org/10.1177/1478077116638921.

Dutton, T. A. (1987). Design and studio pedagogy. *Journal of Architectural Education*, *41*(1), 16–25. https://doi.org/10.1080/10464883.1987.10758461.

Foucault, M. (1977). *Discipline & Punish: The birth of the prison*. Allen Lane. Trans.)A. Sheridan.

Gero, J., & McNeill, T. (1998). An approach to the analysis of design protocols. *Design Studies*, *19*(1 (January 1998)), 21–61. https://doi.org/10.1016/S0142-694X(97)00015-X.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. AldineTransaction.

Goulding, C. (1999). *Grounded theory: some reflections on paradigm, procedures and misconceptions* (Working Paper WP006/99, Issue.

Gül, L. F. (2018). Studying gesture-based interaction on a mobile augmented reality application for co-design activity. *Journal on Multimodal User Interfaces*, *12*, 109–124. https://doi.org/10.1007/s12193-017-0252-0.

Hewitt, M. (2014). Representational forms and modes of conception; an approach to the history of architectural drawing. *Journal of Architectural Education*, *39*(2), 2–9. https://doi.org/10.1080/10464883.1985.10758387.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Hutchins, E., & Klausen, T. (1996). Distributed cognition in an airline cockpit. In Y. Engeström, & D. Middleton (Eds.), *Cognition and communication at work* (pp. 15–34). Cambridge University Press.

Jewitt, C. (2012). *An Introduction to Using Video for Research* (NCRM Working Paper, Issue.

Jewitt, C. (2013). Multimodal methods for researching digital technologies. In S. Price, C. Jewitt, & B. Brown (Eds.), *The SAGE Handbook of Digital Technology Research* (pp. 250–265). SAGE. https://doi.org/10.4135/9781446282229.

Jewitt, C. (2014). *The Routledge Handbook of Multimodal Analysis*. Routledge.

Jewitt, C., & Henriksen, B. (2016). Social Semiotic Multimodality. In N. M. Klug, & H. Stöckl (Eds.), *Handbuch Sprache Im Multimodalen Kontext (Handbook of Language in Multimodal contexts)* (pp. 145–164). De Gruyter.

Jewitt, C., Chubinidze, D., Price, S., Yiannoutsou, N., & Barker, N. (2021). Making sense of digitally remediated touch in virtual reality experiences. *Discourse Context & Media*, *41*(June 2021), 100483. https://doi.org/10.1016/j.dcm.2021.100483.

Klaasen, I. (2002). Modelling reality. In T. M. D. In Jong, & D. J. M. V. D. Voordt (Eds.), *Ways to study and research. Urban, architectural and technical design* (pp. 181–188). DUP Science.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Kress, G., & Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Bloomsbury Academic.

Luck, R. (2014). Seeing architecture in action: Designing, evoking, and depicting space and form in embodied interaction. *International Journal of Design Creativity and Innovation*, *2*(3), 165–181. https://doi.org/10.1080/21650349.2013.875488.

March, S., & Smith, G. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. https://doi.org/10.1016/0167-9236(94)00041-2.

Mewburn, I. (2009). *Constructing bodies: Gesture, speech and representation at work in architectural design studios*. The University of Melbourne].

Mewburn, I. (2012). Lost in translation: Reconsidering reflective practice and design studio pedagogy. *Arts and Humanities in Higher Education*, *11*(4), 363–379. https://doi.org/10.1177/1474022210393912.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, *E77-D*(12), 1321–1329.

Moslehian, A. S., Kocaturk, T., Andrews, F., & Tucker, R. (2022). The nature of innovation in hospital building design: A mixed grounded theory study. *Construction Innovation*, *In Press*. https://doi.org/10.1108/CI-12-2021-0236.

Murphy, K., Ivarsson, J., & Lymer, G. (2012). Embodied reasoning in architectural critique. *Design Studies*, *33*(6), 530–556. https://doi.org/10.1016/j.destud.2012.06.005.

Noghabaei, M., Heydarian, A., Balali, V., & Han, K. (2020). Trend analysis on adoption of virtual and augmented reality in the architecture, engineering, and construction industry. *Data*, *5*, 26. https://doi.org/10.3390/data5010026.

Oh, Y., Ishizaki, S., Gross, M. D., & Do, E. Y. L. (2013). A theoretical framework of design critiquing in architecture studios. *Design Studies*, *34*(3), 302–325. https://doi.org/10.1016/j.destud.2012.08.004.

Oliver, M. (2005). The problem with affordance. *E-Learning and Digital Media*, *2*(4), 402–413. https://doi.org/10.2304/elea.2005.2.4.402.

Picard, R., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D., & Strohecker, C. (2004). Affective learning - a manifesto. *BT Technology Journal*, *22*, 253–269. https://doi.org/10.1023/B:BTTJ.0000047603.37042.33.

Picon, A. (2022). Digital technology and architecture: Towards a symmetrical approach. *Technology | Architecture + Design*, *6*(1), 10–14. https://doi.org/10.1080/24751448.2022.2040297.

Price, S., Jewitt, C., & Yiannoutsou, N. (2021). Conceptualising touch in VR. *Virtual Reality*, *25*, 863–877. https://doi.org/10.1007/s10055-020-00494-y.

Purcell, A. T., & Gero, J. (1998). Drawings and the design process: A review of protocol studies in design and other disciplines and related research in cognitive psychology. *Design Studies*, *19*(4), 389–430. https://doi.org/10.1016/S0142-694X(98)00015-5.

Qureshi, H. (2019). Collaborative architectural design studio environment: An experiment in the studio of Architectural Design-I. *International Journal of Architectural Research*, *14*(2), 303–324. https://doi.org/10.1108/ARCH-12-2018-0049.

Salomon, G., Perkins, D., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, *20*(3), 2–9. https://doi.org/10.3102/0013189X020003002.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (1 ed.). Basic Books.

Schön, D. A. (1985). *The design studio: An exploration of its traditions and potentials* (1 ed.). RIBA.

Sheldon, A., Dobbs, T., Fabbri, A., Gardner, N., Haeusler, M. H., Ramos, C., & Zavoleas, Y. (2019). 15–18 April, 2019). *Putting the AR in (AR)chitecture: Integrating voice recognition and gesture control for Augmented Reality interaction to enhance design practice* 24th Annual Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA 2019): Intelligent and Informed, Victoria University of Wellington, New Zealand. http://papers.cumincad.org/data/works/att/caadria2019_081.pdf.

Sidani, A., Dinis, F. M., Duarte, J., Sanhudo, L., Calvetti, D., Baptista, J. S., Martins, J. P., & Soeiro, A. (2021). Recent tools and techniques of BIM-Based augmented reality: A systematic review. *Journal of Building Engineering*, *42*(October 2021), 102500. https://doi.org/10.1016/j.jobe.2021.102500.

Smithwick, D., & Sass, L. (2014). Embodied design cognition: Action-based formalizations in architectural design. *International Journal of Architectural Computing*, *12*(4), 399–417. https://doi.org/10.1260/1478-0771.12.4.399.

Stals, A., Jancart, S., & Elsen, C. (2021). Parametric modeling tools in small architectural offices: Towards an adapted design process model. *Design Studies*, *72*(January 2021), 100978. https://doi.org/10.1016/j.destud.2020.100978.

Suwa, M., Purcell, T., & Gero, J. (1999). Macroscopic analysis of design processes based on a scheme for coding designer's actions. *Design Studies*, *19*(4), 455–483. https://doi.org/10.1016/S0142-694X(98)00016-7.

Tahsiri, M. (2020). Dialogue in the studio: Supporting comprehension in studio-based architectural design tutorials. *Art Design and Communication in Higher Education*, *19*(2), 149–165. https://doi.org/10.1386/adch_00020_1.

Taneri, B., & Dogan, F. (2021). How to learn to be creative in design: Architecture students' perceptions of design, design process, design learning, and their transformations throughout their education. *Thinking Skills and Creativity*, *39*(March 2021), 100781. https://doi.org/10.1016/j.tsc.2020.100781.

Thompson, J. (2019). *Narratives of architectural education: From student to architect*. Routledge. https://doi.org/10.4324/9781351121873.

Thornberg, R., & Charmaz, K. (2014). Grounded theory and theoretical coding. In U. Flick (Ed.), *The SAGE Handbook of Qualitative Data Analysis* (pp. 153–169). SAGE. https://doi.org/10.4135/9781446282243.

Tsoukas, H. (2017). Don't simplify, complexify: From disjunctive to conjunctive theorizing in organization and management studies. *Journal of Management Studies*, *54*(2), 132–153. https://doi.org/10.1111/joms.12219.

Urquhart, C. (2007). The evolving nature of grounded theory method: The case of the information systems discipline. In A. Bryant, & K. Charmaz (Eds.), *The SAGE handbook of grounded theory* (pp. 339–359). SAGE Publications Ltd. https://doi.org/10.4135/9781848607941.n16.

Urquhart, C., & Fernández, W. (2013). Using grounded theory method in information systems: The researcher as blank slate and other myths. *Journal of Information Technology*, *28*, 224–236. https://doi.org/10.1057/jit.2012.34.

Veliz Reyes, A. (2016). *Augmented pedagogies* [PhD Thesis, University of Liverpool].

Veliz Reyes, A., Kocaturk, T., Medjdoub, B., & Balbo, R. (2012). *Dialogues between physical and digital modelling methods in architectural design* Proceedings of the 30th eCAADe Conference: Digital Physicality | Physical Digitality, Czech Technical University in Prague, Czech Republic. http://papers.cumincad.org/data/works/att/ecaade2012_138.content.pdf.

Verstegen, L., Houkes, W., & Reymen, I. (2019). Configuring collective digital-technology usage in dynamic and complex design practices. *Research Policy*, *48*(6), 103696. https://doi.org/10.1016/j.respol.2018.10.020.

Webster, H. (2008). Architectural education after Schön: Cracks, blurs, boundaries and beyond. *Journal for Education in the Built Environment*, *3*(2), 63–74. https://doi.org/10.11120/jebe.2008.03020063.

Zamani, B., & Babaei, E. (2021). A critical review of grounded theory research in urban planning and design. *Planning Practice & Research*, *36*(1), 77–90. https://doi.org/10.1080/02697459.2020.1830240.