A mixed-methods investigation of how digital immersion affects design for additive manufacturing evaluations

Jayant Mathur
Department of Mechanical Engineering
The Pennsylvania State University
University Park, PA 16802
email: jmm8886@psu.edu

Scarlett R. Miller
Engineering Design and Innovation and Department of Industrial Engineering
The Pennsylvania State University
University Park, PA 16802
email: shm13@psu.edu
Member of ASME

Timothy W. Simpson
Departments of Industrial Engineering and Mechanical Engineering
The Pennsylvania State University
University Park, PA 16802
email: tws8@psu.edu
Fellow of ASME

Nicholas A. Meisel[*]
Engineering Design and Innovation
The Pennsylvania State University
University Park, PA 16802
email: nam20@psu.edu
Member of ASME

ABSTRACT

Applications for additive manufacturing (AM) continue to increase as more industries adopt the technology within their product development processes. There is a growing demand for designers to acquire and hone their design for AM (DfAM) intuition and generate innovative solutions with AM. Resources that promote DfAM intuition, however, historically default to physical or digitally non-immersive modalities. Immersive virtual reality (VR) naturally supports 3D spatial perception and reasoning, suggesting its intuitive role in evaluating geometrically complex designs and fostering

---

[*] Corresponding author

DfAM intuition. However, the effects of immersion on DfAM evaluations are not well-established in the literature. This study contributes to this gap in the literature by examining DfAM evaluations for a variety of designs across modalities using varying degrees of immersion. Specifically, it observes the effects on the outcomes of the DfAM evaluation, the effort required of evaluators, and their engagement with the designs. Findings indicate that the outcomes from DfAM evaluations in immersive and non-immersive modalities are similar without statistically observable differences in the cognitive load experienced during the evaluations. Active engagement with the designs, however, is observed to be significantly different between immersive and non-immersive modalities. By contrast, passive engagement remains similar across the modalities. These findings have interesting implications on how organizations train designers in DfAM, as well as on the role of immersive modalities in design processes. Organizations can provide DfAM resources across different levels of immersion, enabling designers to customize how they acquire DfAM intuition and solve complex engineering problems.

1. INTRODUCTION

Iteration is an essential part of the design process as designers often go through several prototypes of their designs to solve engineering problems. Such prototypes often take the form of sketches, 3D models, or physical props. Physical prototyping, however, can be expensive and time-consuming [1,2], delaying progress to end-use products and solutions. Modern design and manufacturing processes therefore pay special attention to the digital artifact generation that precedes physical fabrication. Organizations leveraging additive manufacturing (AM) to address their end-use product needs are no exception to this. Although AM can reduce the time and cost of physical fabrication, these benefits only materialize when the digital 3D model is favorably designed for AM. The digital 3D modeling and design evaluation stages in the design process are therefore critical. This is because identifying and resolving potential issues with a design early on minimizes build failures and the cost of rework that follows. Designers must therefore know how to evaluate the manufacturability of their designs for the range of potential AM processes. For this purpose, possessing design for AM (DfAM) expertise is a must. However, designers generally lack this expertise, inhibiting them from taking advantage of the fabrication process [3–6]. Practicing DfAM during early design stages fosters the necessary intuition to acquire such expertise [7,8]. Design for AM intuition, therefore, is the designer's ability to evaluate and improve designs for manufacturability by AM by evaluating their opportunistic and restrictive characteristics. Therefore, honing this intuition is essential for designers to innovate with AM and solve complex engineering problems [9].

Possessing DfAM expertise requires distinct instruction on design and process-centric AM concepts, separate from instruction on other manufacturing processes [10,11]. Resources including worksheets [12–15], software tools [16–19], and visualized heuristics [20,21] provide DfAM guidance for this purpose. Because an artifact's design and its evaluations have historically been limited to computer-aided engineering (CAE) tools, designers are habituated to non-immersively evaluate digital designs for manufacturability. As a result, DfAM resources in literature have been modally designed for this established process, yielding digitally non-immersive resources when physical resources are not viable. A recent review of design and manufacturing processes, however, indicates the rise of immersive modalities in design processes for use in 3D modeling, virtual prototyping, and design evaluation [22]. This is because immersive virtual reality (VR) is shown to help designers better perceive the fit, form, and functionality of a design [23,24]. Such enhanced perception also improves the ability to identify errors and defects with 3D models [25,26]. With the advent of more affordable consumer-grade VR headsets, designers can now feasibly leverage the benefits of *immersiveness* in their design processes. However, little work examines varying presentations of DfAM knowledge, including presentations in digitally immersive modalities [27,28]. The benefits of immersion specifically on DfAM evaluations are therefore not well understood. There is a need to investigate how immersive DfAM evaluation affects the outcomes and effort associated with the act of evaluating designs. This research addresses this gap in the literature by leveraging established DfAM resources in immersive modalities and investigating the effects of immersion on DfAM evaluation.

Additive manufacturing encourages solutions with unique geometric complexity that are difficult to achieve with other manufacturing processes. This complexity often takes the form of organic, generative, or lattice structures that are uncommon in designs for subtractive manufacturing processes. Incorporating such complexity can be instrumental to the desired solution, but it can make the digital design difficult to evaluate for manufacturability. Given the 3D nature of geometric complexity, leveraging 3D spatial immersion in VR to aid DfAM evaluations seems more intuitive than using non-immersive CAE. However, before organizations use VR for this purpose research must establish how digitally immersive DfAM evaluations compare to their digitally non-immersive experiences and physical counterparts. This requires an investigation into how varying levels of immersion affect DfAM evaluation processes, which is currently lacking in the literature. The goal of this work is to, therefore, investigate the design of VR experiences for DfAM evaluations. For this purpose, Section 2 reviews the literature, presenting the relevant guidance behind the proposed research scope in Section 3. Section 4 describes the method of study used to address the research questions and Section 5 presents the results from the study. Additional details on the findings and their implications are then discussed in Section 6, with Section 7 summarizing the collective contributions of this work and its limitations for future work. The contributions of this work have significant implications for how future designers are trained in DfAM to meet the AM demands in the workforce.

2. LITERATURE REVIEW

Before investigating the effects of immersion on DfAM evaluations, it is important to clarify the meaning of immersion. Digital immersion and presence refer to how well the environment can mimic visual, auditory, and other sensory aspects of the physical world [29–31]. They determine how compelling, engaging, and educationally meaningful VR experiences are perceived [32–34]. As a result, traditional computer displays generally fall under non-immersive VR while head-mounted displays (HMDs) fall under immersive VR. For the scope of this research, immersion and presence are collectively referenced as *immersion* along with the following distinctions: CAE = non-immersive virtual modality (i.e., a flat-screen computer display), VR = immersive virtual modality (i.e. an HMD with controllers), REAL = immersive physical modality (i.e. the physical world). The remainder of this section reviews past literature to support exploring DfAM experiences in VR with these distinctions in mind. First, Section 2.1 compares and contrasts immersive and non-immersive modalities to highlight the effects of enhanced digital immersion on user experiences. Next, Section 2.2 identifies the gap in AM literature regarding immersive DfAM experiences and motivates the need to design and study such experiences.

2.1 Effects of digital immersion on user experiences

Organizations have historically leveraged non-immersive computer-aided experiences to replicate in-person experiences [35]. Computer-driven, game-based [35–37] or simulation-based [38–41] design and manufacturing experiences have garnered key attention for this purpose. This is because, like in-person experiences, they foster technical and professional skills by challenging users to reflect on the impact of their decisions [38]. They also improve collaborative learning in problem-solving situations [42] and promote high states of concentration and satisfaction at low cognitive loads [37]. Enhanced spatial immersion in modalities like VR adds to CAE experiences by enabling improved design conceptualization and analysis through enhanced spatial immersion [43]. Compared to REAL and CAE modalities, VR engagement can bolster design creativity [44] and concept generation [45]. Additionally, engaging in VR increases enjoyment, learning outcomes [46], and self-efficacy [47] when compared to passive and non-immersive video-based engagement. Furthermore, research indicates improved acquisition of declarative and procedural knowledge [29] and cognitive and affective skills [30] when working in VR. This further influences memory recall, affecting the application of such knowledge and skills [48].

Past work supports using immersive modalities for DfAM training by emphasizing their benefits over non-immersive modalities. However, the overall effects of working in VR are not entirely understood and continue to be investigated. The literature emphasizes a strong influence of environmental factors on the effects of VR engagement [49–53], specifically the influence of environmental and pedagogical conditions of the designed immersive experiences on meaningful outcomes [54–57]. The observed cognitive load is similarly influenced by the manual operations required by the environment during design and learning experiences [58–61]. This means that the effects of immersion on user experiences are not universal and are instead influenced by the context of the experience. This breadth of evidence demands an investigation into the effects of the modality in DfAM contexts. First, research must examine the effects on spatial perception and reasoning and the user's psychomotor engagement [62]. Second, an investigation into the effects of the experienced cognitive load on the experienced learning must be conducted [63,64]. These goals are essential to understanding how DfAM experiences in VR vary from other experiences.

2.2 The lack of digitally immersive DfAM resources

Designers must be trained in considering design for manufacturing (DfM) and visually checking their design's manufacturability. Best suited for early-stage problem-solving, such visual checks hone a designer's tacit design knowledge to promote innovation and increase design quality [11,65] while minimizing development time and cost from rework [66]. Additionally, acquiring such tacit knowledge from visual DfM evaluations better prepares designers to use automated analysis tools [16–18] for end-design stages where the cost of rework is high. The importance of DfM considerations in design processes has historically yielded several resources that help designers acquire and apply DfM intuition. However, DfAM resources are not widely incorporated alongside these DfM resources. Modern DfM tools help foster subtractive design thinking that is required for other manufacturing processes [11,67]. There is a similar need for DfAM resources that specifically support additive, generative, and organic design thinking that is required for AM processes [10,11,68,69].

Although literature offers different worksheets [12–15], frameworks [19,70], and heuristics [20,21] for this purpose, they are limited to comprehension within a physical or digitally non-immersive modality. Limited research compares modalities on their presentation of DfAM concepts [27,28], with only previous work by the authors further exploring the effects of immersion on DfAM evaluations [28]. There is a need to strengthen our understanding of how differences in immersion between modalities affect DfAM consideration during design evaluations. Research indicates that adding VR immersion promotes improved capabilities in manufacturing and assembly conditions [71,72]. Added immersion and presence strongly influence the perception of 3D designs [73] and other presented information [74]. Compared to CAE evaluations, this enhanced perception improves the ability to identify errors and defects with 3D models [25,26] and better perceive the dimensional fit of a design [23,24]. Perception of 3D designs is a key factor that affects visual DfAM evaluations and is uniquely different in VR than it is in CAE or REAL conditions. Just as sketching is suited for highlighting design concepts while CAD is better equipped for detailed engineering [75], research must evaluate the role of VR in DfAM processes. The goal of this research is to contribute to this gap in AM literature by investigating the effects of immersion on DfAM evaluations. The presented work specifically identifies how VR evaluations compare to CAE and REAL evaluations.

3. RESEARCH QUESTIONS

This research aims to investigate how immersion affects experiences involving the DfAM evaluation of 3D models. For this purpose, this research implements a mixed methods study with a sequential explanatory design. In other words, the study first extracts quantitative information and then qualitative information on the observed effects. Such an investigation offers a general understanding of the trends that exist as well as insight into the underlying mechanism behind these trends. The following research questions (RQs) guide this investigation:

Research Question 1. *How do the differences in immersion between CAE, VR, and REAL modalities affect the outcomes of DfAM evaluations of designs of varying manufacturability?*

This research question identifies the effects of immersion on DfAM evaluation outcomes by examining the trends in quantitative data. Specifically examined are 1. the DfAM score of the design, 2. the time taken for the evaluation, and 3. the confidence of the evaluation. Compared to the CAE evaluation, it is hypothesized that the VR and REAL evaluations will yield scores closer to expert scores from faster and more confident evaluations. However, no significant differences between the two immersive modalities are expected. Such trends are hypothesized due to expected enhancements in spatial perception and reasoning within immersive modalities [23–26]. Effects on DfAM reasoning from the perceived complexity of the evaluated designs are also expected. Specifically, the difference in outcomes between the immersive and non-immersive evaluations is expected to increase for designs with higher perceived complexity [23,24].

Research Question 2. *How do the differences in immersion between CAE, VR, and REAL modalities affect the cognitive load experienced when evaluating designs of varying manufacturability?*

This research question extracts further quantitative insight into the effects of immersion on completing DfAM evaluations. Specifically examined is the numeric, self-reported, cognitive load experienced by the designers during the evaluation. Unlike RQ 1, this research question focuses on the effort required to complete the *entire* DfAM evaluation exercise. This approach is synonymous with design processes where designers must make conclusions about the design from iterative evaluations using the same modality. Compared to the CAE experience, it is hypothesized that the VR and REAL experiences will generally yield lower reported cognitive load values. However, no significant differences between the two immersive modalities are expected. It is expected that the effort required to perform design evaluation operations, and thus the cognitive load, changes due to the change in immersion. Specifically, evaluations within modalities that require low effort will yield lower reported cognitive load than those that require high effort [58,60,61,76]. Such variation in effort is expected to arise due to differences in immersion, the perceived complexity of the designs, and the required engagement with the designs.

Research Question 3. *How do the differences in engagement between CAE, VR, and REAL modalities explain the observed trends in DfAM evaluation outcomes and cognitive load?*

This research question dives into the mechanics of completing DfAM evaluations within varying levels of immersion by examining qualitative data. Specifically, the designer's active and passive engagement with the designs during the DfAM evaluation is observed. Given that designers are studied in similar environments, it is hypothesized that analyzing how they engage with the designs will explain the trends observed in RQ 1 and RQ 2. This is expected because when other factors are controlled, the differences between groups can likely be attributed to their modality. Specifically, attributed to how immersion alters the perception of 3D artifacts and other visual information [73,74] and influences the interactions involved. As an explanatory research question, the goal is to establish a basic understanding of how designers interact with and evaluate designs for AM within varying levels of immersion. Such insight lays the groundwork for future hypothesis-driven research into utilizing different modalities for DfAM problem-solving.

4. METHODOLOGY

This research was motivated to observe the role of immersion in cultivating a designer's DfAM intuition under given printing constraints. Specifically, to understand how immersion affects a designer's evaluation of different designs for their manufacturability with AM at specific print orientations and process parameters. The goal of this research was to, therefore, identify the effects of immersion on DfAM evaluation when evaluating designs of varying manufacturability.

Designers assessed manufacturability for material extrusion AM through either the CAE, VR, or REAL modality. Although print orientation and process parameters affect a design's manufacturability, through outcomes such as print time, support usage, etc, this research focused on studying circumstances that resemble a visual check of the design *before* calculating such manufacturability outcomes. Such an investigation informed on the role of immersion in the designer's reasoning and thinking process during DfAM evaluations. A mixed-methods study of their experiences was conducted using a sequential explanatory design, i.e., a quantitative phase followed by a qualitative phase. The design of the experiments for both phases was similar, except for a think-aloud task included in the qualitative phase during the DfAM evaluation exercise. This think-aloud data from every participant's perspective explained their engagement with the designs and contextualized the general trends observed in the quantitative results.

The designed experiment required completing the steps illustrated in Fig. 1. All these steps were completed on an online Qualtrics survey. This includes presenting questionnaires, instructions, and the CAE and VR digital tools for the DfAM evaluations. No personal or identifiable information was collected from the participants. Completing the survey corresponded to opting in for the study; not doing so was registered as *opting out* of the study. Participant data was deleted accordingly as per the approved Internal Review Board (IRB) protocol.

Section 4.1 describes the process of using the survey to assign participants to the study conditions and ask about their backgrounds. Specifically, their background in AM, material extrusion (ME), design for ME (DfME), and proficiencies with CAE and VR. These were all recorded on a 5-point Likert scale. After sharing their backgrounds, participants were introduced to their assigned modality and a tutorial. The goal of this tutorial was to provide familiarity with the modality and the DfAM evaluation exercise. Doing so helped minimize the effects of technological proficiency on the measured outcomes in the study.

Participants assigned DfAM scores to one design during the tutorial and three designs during the main study. These designs came from a set of six pre-selected 3D models. Section 4.2 shows this set of designs and explains the expert review process used to select them. For the main study, each design was evaluated one at a time, in a pre-determined counterbalanced order as further explained in Section 4.3. For this DfAM exercise, each evaluation was measured on three outcomes: 1. the design's DfAM score, 2. the time taken for the evaluation, and 3. the confidence of the evaluation. The DfAM score was the calculated sum of eight distinct process-agnostic metrics. These metrics were consolidated from past work [12,13] into a worksheet [77] that was provided for DfAM evaluations. Each metric was evaluated on a 3-point Likert scale, corresponding to a *low-medium-high* scale. Designs therefore received a DfAM score between 8-24 points, with a higher score suggesting higher manufacturability with ME. Participants in the qualitative method group were additionally asked to think aloud during the DfAM exercise (see Section 4.4). Only three designs were presented to minimize the effects of survey fatigue. Doing so retained focus on the cognitive load directly impacted by the exercise of evaluating designs within their assigned modality.

Upon completing evaluations for three designs, participants reported the cognitive load they experienced from the exercise. Section 4.5 explains how this data was collected with a self-reported Workload Profile Assessment (WPA) [78]. The tool measured the cognitive load exerted during the experience across eight dimensions. Each dimension was scored between 0 and 10 to represent their cognitive load. This data offered context to the effort required for completing the DfAM evaluation exercise within each modality. Pairing this information with that derived from the DfAM exercise data holistically demonstrates how varying levels of immersion affect DfAM processes.

4.1 Pre-study procedure

Participants in this research were second and third-year undergraduate students. They were recruited from an engineering design methodology course at an R1 university. All students were informed of their rights and options as per IRB protocol before conducting the study. Those who *opted in* to participate were given an online Qualtrics survey to use for their participation in the study. Hidden from the participants, the survey's built-in algorithm balanced assignments evenly between the three conditions: CAE, VR, or REAL. Once assigned to a condition, participants answered a questionnaire, describing their knowledge of 3D printing. This data encapsulated their general experiences with AM and their specific experiences with the ME process and DfME practices. Collectively, this data indicated the need to account for prior knowledge in the statistical analysis of the measured outcomes in the study. Knowledge in AM, ME, and DfME was recorded on the following 5-point Likert scale:

1. I have never heard or learned about this topic before this
2. I have some informal knowledge on this topic
3. I have received some formal knowledge on this topic
4. I have received lots of formal knowledge on this topic
5. I am an expert on this topic

After sharing their prior knowledge of 3D printing, participants in the CAE and VR conditions described their proficiency with their assigned modality. Those in the REAL condition were not asked for any such proficiency. As relevant to the study, the questionnaire specifically inquired about their proficiency in working with or interacting with 3D models. This data served to establish the need for a tutorial phase for each condition before the main study. This is because the participants likely had much more experience working in CAE and REAL modalities than in VR. An on-par comparison of DfAM processes between the conditions necessitated a tutorial, requiring empirical evidence for support. Therefore, it was important to measure and acknowledge this difference in technological proficiency. Proficiency in CAE and VR was recorded on the following 5-point Likert scale:

1. I have never worked with 3D models in this modality before this
2. I am slightly comfortable working with 3D models in this modality
3. I am comfortable working with 3D models in this modality
4. I am extremely comfortable working with 3D models in this modality
5. I am an expert on working with 3D models in this modality

Upon completing the questionnaire, participants were introduced to their assigned modality. Those assigned to the CAE condition were directed to the evaluation activity via a link on their computers. Those in the VR and REAL conditions were directed to designated areas that were set up with the resources required for their respective conditions. Participants in the VR condition were each given a Meta Quest headset and controllers and directed to the evaluation activity on the Meta Quest Browser app. Participants in the REAL condition were directed to a table with the physical parts where they continued the survey. The physical parts were manufactured using ME and underwent multiple post-processing cycles of coating with primer and sanding. Doing so minimized any visible indications of the original fabrication process, minimizing biased evaluations. Once at their designated areas, participants proceeded to the tutorial: a practice DfAM evaluation exercise designed to familiarize them with their assigned modality.

4.2 Selecting 3D models

The goal of the design selection process was to identify a set of designs that *truly* varied in their DfAM scores. This is because this research aimed to identify the effects of immersion on DfAM evaluation when evaluating designs of varying manufacturability. Studying designs with identical DfAM scores could inhibit isolating the effects of immersion on the measured outcomes [28]. For this purpose, this research first identified a set of designs to use for the main study through an expert review process. Six experts with 4-10 years of demonstrated AM and DfAM expertise in academia and industry reviewed twelve different designs pre-selected by the authors. The experts carried out the same DfAM evaluation exercise prepared for the main study for each design. This means each design was evaluated using eight metrics on a 3-point Likert scale [77]. The sum of these was an expert-established DfAM score, with a higher score indicating higher manufacturability.

To optimize the sample size for statistical analysis, only six of the designs from the original twelve were selected for the main study (see Fig. 2 and Ref. [79]). For this selection, the designs were roughly grouped into low, medium, and high-scoring designs. *Lows* were scored roughly between 8-13, *mediums* between 14-18, and *highs* between 19-24. Two designs from each group were selected, specifically, those that varied the most in their DfAM scores between the groups. There was a significant difference in DfAM scores between the low group (D1, D2) and the high group (D5, D6). The differences from the medium group (D3, D4) were not as significant but still observable and therefore included in the main study.

4.3 The DfAM exercise

The goal of this research was to observe how varying levels of immersion affect DfAM evaluations. As a result, participants were tasked with evaluating 3D models for manufacturability in either the CAE, VR, or REAL modality. Each condition included three key features to aid the DfAM evaluation: 1. virtual or physical 3D models, 2. tools for measuring and evaluating the designs, and 3. digital instruction on completing the exercise. To ensure that the exercise was similar across the conditions, all the digital and physical features were made identical. Figure 3 presents the designed environments for each condition to demonstrate this.

As shown in Fig. 3, participants were instructed identically to evaluate the designs for manufacturability in a pre-defined, but not necessarily optimal, print orientation. They were reminded to consider the ME process during the design's evaluation. Free interaction with a design and its environment was permitted to encourage intuitive exploration of the designs. As such, typical engagement in VR and REAL included picking up, rotating, and moving the models to get a good view of the design. Those in CAE manipulated the camera by zooming, orbitally rotating, and panning for the same purpose. Each modality afforded interactions that compensated for its inherent limitations, enabling similar experiences across the modalities. Designs were also presented at a fixed scale in all the modalities. Therefore, rescaling or digitally enlarging the 3D model in CAE and VR was not permitted. This means that the dimensions of the digital models matched those of the physical objects, ensuring identical comparisons between the modalities.

While evaluating designs on DfAM, participants used a worksheet [77] with eight metrics derived from past work by Booth et al. [12] and Bracken et al. [13]. These metrics corresponded to the following eight DfAM concepts:

1. Removal of support structures
2. Presence of unsupported overhangs
3. Presence of unsupported bridges
4. Presence of self-supporting features
5. Sharpness/Rounding of cross-sections
6. Size/Area of cross-sections
7. Thinness of features compared to the print resolution
8. Surface finish on non-build direction curved surfaces

Each metric was evaluated on a 3-point Likert scale, resulting in a sum score of 8-24 points for each design. For the main study, participants were not informed of these DfAM scores. However, during the tutorial, they were offered a comparison of their evaluation with an expert's. This means that participants were treated like experts in the main study and were not provided with any feedback on their evaluation.

Participants concluded one evaluation by filling out the entire worksheet and reporting their confidence in the design's evaluation. They completed the entire exercise by evaluating three designs, one at a time. Limiting evaluations to three designs minimized the effects of survey fatigue. This retained a focus on studying the cognitive load experienced directly from completing the DfAM exercise within their assigned condition. The three designs presented were arbitrarily assigned from six possible options (see Section 4.2). A 6x6 Balanced Latin Square was generated and split into two 6x3 tables, presenting 12 distinct orders to use for the study. These orders were counterbalanced, thus, minimizing immediate sequential or carry-over effects [80].

4.4 Think-aloud protocol

The goal of the think-aloud task was to understand the participants' engagement with the designs during the DfAM evaluation exercise. For this purpose, the following think-aloud protocol was implemented for the experiment:

- The task was untimed and participants were encouraged to take their time evaluating the designs.
- Participants were prompted to explain *"What about the design rationalized the option(s) you[participant] chose on the DfAM worksheet?"*.

- They were instructed to verbalize all their thoughts as frequently as possible.
- If they were silent for more than 30 seconds, they were reminded to *think aloud* and continue.

To collect think-aloud data, each participant's session was video and audio-recorded. Participants were informed of this recording and were asked to re-confirm their consent before proceeding. Those who changed their consent were not recorded and were excluded from the study. The video and audio recordings were later coded together to extract the think-aloud data. Two expert raters (from the authors) coded all the recordings using these codes in DARMA, a joystick-driven, dual-axis rating tool for videos [81]. To clarify coded versus uncoded content, raters moved the joystick to the extremes of the axes when assigning codes to the recordings. Doing so ensured that the coded data was distinguishable from the uncoded data. The final two-dimensionally coded data was used to explain the participants' engagement with the designs during the DfAM evaluation exercise.

A deductive coding approach was used to analyze the think-aloud data. Themes were deduced based on work by Lauff et al., which informs on how designers engage with artifacts in design processes [82]. Specifically, how designers actively and passively engage with artifacts. These themes expand upon their roots in design communication [83,84] and are similarly utilized in past work with 3D artifacts and VR contexts [85,86]. Based on these thematic distinctions, *Engagement* in this work was defined as *active* and *passive* interaction with the designs. Active engagement corresponded to *direct* interactions with the 3D objects. This included picking up, rotating, and moving the models to get a good view of the design. Passive engagement corresponded to *indirect* interactions with the 3D objects or making contextual references to the designs. Pointing at the design and its features without *intentionally manipulating* the model was considered indirect interaction. Emphasizing aspects of the design that were related to AM or DfAM concepts was considered contextual referencing.

The two codes established with these themes in mind were *Referencing* and *Interacting* (see codebook in Table 1). These corresponded to *passive* and *active* engagement with the designs respectively. In DARMA, the code *Interacting* was assigned to the axial ends of the x-axis to signify active engagement. Similarly, *Referencing* at the ends of the y-axis signified passive engagement. Axial polarity was irrelevant to the coding process and codes were standardized to the same quadrant for analysis. The center of the axes was assigned as *No Engagement*. This was used to signify the absence of any engagement with the designs and account for time spent on other unrelated tasks. The coding process was on a continuous timeline, meaning recordings were not segmented or discretized. The codes were not mutually exclusive and were assigned to the same time point if appropriate. This means that every time point in the recording corresponded to *No Engagement*, *Referencing*, and/or *Interacting*, and no point was left uncoded.

4.5 Gauging cognitive load

After completing the design evaluation exercise, participants reported their cognitive load. They used the Workload Profile Assessment (WPA) tool [78] to quantify the cognitive load they experienced. Compared to the Subjective Workload Assessment Technique and the NASA Task Load Index, the WPA's higher sensitivity was preferred for such quantitative assessments [87]. Participants scored eight workload profile dimensions between 0 and 10 to represent their mental exertion. These eight dimensions spanned Perceptual, Response, Spatial, Verbal, Visual, Auditory, Manual, and Speech cognitive processing needs. Participants received a text and audio description of each dimension to review, along with an example of each dimension applied in practice.

Using these descriptions, participants assessed their cognitive load and assigned appropriate values to each dimension, one at a time. The *Verbal* and *Auditory* dimensions, though not directly applicable to the design DfAM exercise, were included to ensure consistency with the WPA tool. This is because the designed experiment did not study any tasks or elements that gave verbal instruction and audio cues. However, the WPA tool was designed to study tasks that would include such elements. The *Speech* dimension was also included under the same rationale. Although the think-aloud task in the qualitative method group induces *Speech* processing cognitive load, this research limits measurements of mental exertion to the quantitative method phase. Additionally, this research did not check or correct for any misinterpretations of the dimensions by the participants. Therefore, the inclusion of these dimensions ensured that all the necessary information was collected to study the cognitive load experienced by the participants.

5. RESULTS

This research conducted a mixed-methods study to evaluate the effects of immersion on DfAM evaluation and experiential cognitive load. To statistically explain the background data and the cognitive load data, linear regression models (lm) were generated. Linear mixed-effects regression modeling (lmer) was used to statistically analyze the DfAM evaluation data. Pairwise comparisons between variables were done using Estimated Marginal Means tests. The *lmer* utilized restricted maximum likelihood estimation to iteratively modify the parameter estimates with a minimized log-likelihood function. The *lm* and *lmer* model assumptions were checked using the Peña and Slate [88] and the Loy and Hofmann [89] procedures respectively. Unless otherwise specified, this research did not find any observable violations and relies on the acceptable range for the robustness of the respective regression models. A 95% confidence interval was used to determine statistical significance (i.e., $p < 0.05$). The p-values from the *lmers* are adjusted using the Kenward and Rogers adjustment to account for the small sample size. Those from the pairwise comparisons were adjusted using the Bonferroni method to account for multiple comparisons. All potential outliers in the data were retained in each analysis. The reported findings are presented in the following format: $b = 0.00, F(n,m) = 0.00$ $[t(n,m)$

= 0.00], p = 0.00. Here, *b* is the regression coefficient (i.e., slope), *F* is the F-statistic, *t* is the t-statistic, and *p* is the p-value. The *n* and *m* values represent the degrees of freedom for the numerator and denominator respectively.

5.1 Background analysis

The study included 124 participants between two method studies: Quantitative and Qualitative. As shown in Table 2, they were evenly distributed in each method across the three conditions: CAE, VR, and REAL. Note that Table 2 lists only the participants who completed *all* the required tasks of the study for their method group. Also, note that only participants in the qualitative method group conducted the think-aloud task during the DfAM exercise. The distribution in Table 2 was uniform within acceptable margins, strengthening the statistical analysis of the measured outcomes.

Analyzing the participants' prior knowledge of AM, ME, and DfME concepts helped account for the effects of such knowledge on the measured DfAM outcomes and cognitive load. For the analysis, the distributions of the prior knowledge in AM, ME, and DfME were regressed on the centered condition (CAE= -0.5, VR= 0, REAL= 0.5). The results showed no observable significant difference between the three conditions in their prior knowledge of AM, $b = -0.09$, $F(1,122) = 0.26$, $[t(1,122) = -0.51]$, $p = 0.612$, of ME, $b = -0.19$, $F(1,122) = 0.78$, $[t(1,122) = -0.88]$, $p = 0.38$, and of DfME, $b = -0.15$, $F(1,122) = 0.47$, $[t(1,122) = -0.69]$, $p = 0.493$. This trend can be observed in Fig. 4, where participants in all the conditions reported similar prior knowledge of AM, ME, and DfME. Specifically, they shared that they generally had *some informal knowledge* of each of the topics.

Participants in the CAE and VR conditions also described their proficiency with their respective modalities. Analyzing this data established the need for a tutorial phase for each condition before the main study. The collapsed technology proficiency was regressed on the centered condition (CAE= -0.5, VR= 0.5). As expected, participants generally showed a significantly higher proficiency for CAE technology than for VR technology, $b = -1.66$, $F(1,83) = 45.74$, $[t(1,83) = -6.76]$, $p < 0.001$. Specifically, participants in the CAE condition were generally *extremely comfortable* with CAE technology; however, those in the VR condition had generally *never worked with* VR technology. This trend shown in Fig. 5 was expected because students had likely completed CAE/CAD course requirements but likely not any VR coursework. Although expected, the trend supports the need for a tutorial on working in VR. Such a tutorial balances the technological proficiency between the conditions before the main DfAM study.

5.2 DfAM outcomes

Results presented in this section are observations of the quantitative method group only. This group did not conduct the think-aloud task during the DfAM exercise and consisted of only 93 participants. Analyzing the data collected from this group helped tackle the first research question, i.e., identifying the effects of varying levels of immersion on DfAM outcomes. For this analysis, the DfAM score, evaluation time, and reported confidence were regressed on the centered variables for *Condition*, and *Design* as a covariate. *Condition* served as a between-subjects variable centered around the three studied conditions: CAE = -0.5, VR = 0, REAL = 0.5. *Design* served also as a within-subjects variable centered around the six designs: D1 = -0.5, D2 = -0.3, D3 = -0.1, D4 = 0.1, D5 = 0.3, D6 = 0.5. Each participant's unique ID (*PID*) served as a random intercept to control for non-independence of observations. The presented results from the regression analysis focus on each detailed effect when controlling for all other main effects in the model. Only the interaction effects between condition and design were considered in the analysis.

The main analysis showed no significant effect of *Condition* on the *Score*, *Time*, and *Confidence* (see Table 3a). As seen in Figures 7, 8, and 9, participants generally reported similar scores, experienced similar evaluation times, and were equivalently confident across the modalities. The main analysis also showed no significant effect of design on *Time* and *Confidence* but showed a significant effect on the DfAM scores (see Table 3b). Specifically, on collapsing *Condition*, participants reported significantly higher DfAM scores as the designs changed from D1 to D6. Figures 7, 8, and 9 show that participants identified significant differences between the designs themselves regarding their manufacturability by ME, with similar amounts of time and confidence. This means that the selected designs were suggestive of varying DfAM scores and that participants could intuit this.

Estimating a two-way interaction between *Condition* and *Design* explained how the effects of the modalities on the DfAM outcomes varied with the designs. The analysis showed a significant effect from the interaction between *Condition* and *Design* on *Score*, but not on *Time*, and *Confidence* (see Table 3c). Specifically, the effect of *Condition* on *Score* decreased as the value of *Design* changed from D1 to D6. Figure 6 illustrates this interaction effect where the direction of the effect of *Condition* on *Score* flips from D4 to D6. In other words, the DfAM scores from the CAE condition go from being lower than the VR and REAL scores to being higher than them. This means that a modality's effect on the DfAM scores was dependent on the design being evaluated.

A secondary analysis was conducted to understand the effects of *Condition* and *Design* on the difference between the DfAM scores assigned by participants and the expert scores. This analysis showed no significant effect of *Condition* on the difference between scores, $b = 0.39, F(1,91) = 2.09, [t(1,91) = 1.45], p = 0.152$, but showed a significant effect of *Design* on the difference, $b = 0.64, F(1,267) = 4.79, [t(1,267) = 2.19], p = 0.029$. Specifically, on collapsing *Condition*, DfAM scores assigned by participants generally deviated further from the expert scores as the designs changed from D1 to D6. This means that although participants could identify differences in the designs' manufacturability, they could not evaluate them as well as the experts. That is participants under or over-estimated the expert scores by similar amounts across the conditions, which worsened as the designs changed. The estimated two-way interaction between *Condition* and *Design* showed no significant effect on the difference between scores, $b = 0.3, F(1,263) = 0.17, [t(1,263) = 0.41], p = 0.679$. This means that the effect of *Condition* on the difference between scores was not dependent on the design being evaluated.

5.3 Cognitive load

Results presented in this section are also observations of the quantitative method group only (i.e., from 93 participants). Analyzing the data collected from this group helped tackle the second research question, i.e., identifying the effects of varying levels of immersion on cognitive load. Checking the assumptions for linear regression modeling showed violations of normality in the data for the *Auditory* and *Speech* dimensions. These are sensible violations because the *Auditory* and *Speech* dimensions did not apply to the DfAM exercise. In addition to the *Verbal* dimension, these dimensions were excluded from the analysis and the reported findings. For this analysis, the *remaining* five dimensions of cognitive load were regressed on the centered variable for *Condition*. *Condition* served as a between-subjects variable centered around the three studied conditions: CAE = -0.5, VR = 0, REAL = 0.5.

The main analysis showed no statistically significant difference in cognitive load for any of the dimensions across the three conditions (see Table 4). As observed from Fig. 10, this suggests all the conditions demanded similar effort in processing evaluations across the different dimensions from the participants. Although there were generally no significant effects on each dimension, an emerging trend for the *Visual* dimension can be observed. This trend is seemingly driven by the immersive conditions. Specifically, participants reported lower *Visual* cognitive load as the condition changed from CAE to VR to REAL. However, the standard deviation of the current dataset inhibits acquiring concrete information on the trend.

5.4 Modality engagement

Results presented in this section are observations of the qualitative method group only. This group conducted the think-aloud task during the DfAM exercise and consisted of only 31 participants. The video and audio recording of each participant's session was coded for the think-aloud task to extract the data. Of the 31 recordings, 6 recordings (two per condition) were randomly selected to establish reliability between the two raters. The remaining 25 recordings were divided between the two raters and coded independently. Reliability between raters was established using the Intraclass Correlation Coefficient (ICC) for average and consistent agreement (see Table 5). Any disagreements between the raters were resolved through discussion and consensus.

The data collected from the think-aloud task was analyzed to understand the differences in engagement between the modalities. For this analysis, the engagement was regressed on the centered variable for *Condition* independently for each type of engagement: *Active* and *Passive*. *Condition* served as a between-subjects variable centered around the three studied conditions: CAE = -0.5, VR = 0, REAL = 0.5. Engagement was equated to the number of codes observed per minute of the recording, i.e., the ratio of the total number of codes to the total recording time.

The main analysis showed a significant effect of *Condition* on *Active* engagement, *b = -5.33, F(1,73) = 6.65, [t(1,73) = -2.58], p = 0.012*, but showed no significant effect on *Passive* engagement, *b = -0.33, F(1,73) = 0.03, [t(1,73) = -0.17], p = 0.866*. Specifically, participants generally demonstrated higher *Active* engagement in CAE than in VR and REAL, while *Passive* engagement was similar across the conditions (see Fig. 11). This means that participants generally manipulated the designs more in CAE than in VR and REAL during their DfAM evaluations. Figure 11 further shows that the pairwise comparisons of *Active* engagement were significant between CAE and VR, and CAE and REAL, but not between VR and REAL. This means that the differences in *Active* engagement were significant between immersive and non-immersive modalities, but not between the immersive modalities themselves.

6. DISCUSSION

Results suggest that outcomes from the DfAM evaluations do not observably vary with the immersion level of the modality. However, the relationship between the immersion level and the outcomes was found to be dependent on the design being evaluated. Additionally, the cognitive load experienced from the evaluations does not vary with immersion; however, emerging trends were observed. Furthermore, the results suggest that a designer's passive engagement with designs does not vary with immersion, but their active engagement does.

The observed findings suggest interesting implications for research in immersive DfAM experiences and the development of digital design experiences with AM. First, these findings demonstrate the potential for immersive VR as a complementary resource to CAE and REAL DfAM decision-making. In other words, designers can transition between CAE, VR, and REAL modalities as preferred without significantly affecting their DfAM outcomes or cognitive load. Such flexibility may also apply to broader DfM workflows that strategically leverage AM with other manufacturing processes. Second, the findings also inform the design of learning modules, indicating the potential for instructors to strategically incorporate VR to intuitively instruct certain DfAM and DfM concepts. As shown in Fig. 12, VR can be significantly more enjoyable than CAE and REAL for novice users to cultivate DfAM and DfM intuition, with little to no effect on the experiential outcomes. Beyond such broad implications, the remainder of this section breaks down these findings and their implications as they specifically relate to the research questions.

How do the differences in immersion between CAE, VR, and REAL modalities affect the outcomes of DfAM evaluations of designs of varying manufacturability?

The goal of RQ 1 was to understand how differences in immersion between modalities affect the outcomes from the DfAM evaluations. It was hypothesized that the VR and REAL evaluations would yield scores closer to expert scores from faster and more confident evaluations, as compared with the CAE evaluations. No significant differences between VR and REAL were expected. Regarding the primary effects of immersion, the results in Section 5.2 failed to reject the null hypothesis for each outcome. Specifically, the study could not identify significant differences in DfAM score, evaluation time, and reported confidence between the conditions. The offsets of participant scores from the expert scores were also not significant. In other words, participants consistently under or over-estimated the expert scores by similar amounts across the conditions. It was also hypothesized for RQ 1 that the difference in outcomes between the immersive and non-immersive evaluations would increase for designs with higher perceived complexity. In other words, the effect of immersion on the outcomes was expected to be dependent on the design being evaluated. The results in Section 5.2 showed a significant interaction effect between *Condition* and *Design* on the DfAM scores. However, contrary to the hypothesis, the effect of immersion on the DfAM scores decreased as the designs changed from D1 to D6. Specifically, the DfAM scores from the immersive conditions went from being higher than the CAE scores to being lower than them. Comparing these findings to the results in Table 3b indicates that participants seemed to lean toward a *neutral* evaluation as the designs changed from D1 to D6.

These are interesting findings because they imply that the modality for evaluating 3D artifacts may not be the driving factor for manufacturability-by-AM evaluations. That is, invoking changes to the mental models of novice designers regarding their application of DfAM may not be driven by digital or physical immersion. Of particular interest here are the observations regarding the REAL condition. Specifically, the implication that physical artifacts may not be necessary for manufacturability-by-AM evaluations, and digital artifacts may suffice. As it stands, participants seem adept at identifying unfavorable and favorable features in the designs but fall short of evaluating them *expertly*. It is worth noting that not identifying such distinctions could have inhibited the study's ability to identify the effects of immersion on DfAM outcomes [28]. Specifically, the lack of diversity in the designs would yield similar scores across the designs, masking any observable differences between the conditions. Therefore, the observed findings may be attributed to either the participants' lack of expertise in AM and DfAM or the nature of the DfAM worksheet used in the study. If the former is valid, a designer's established expertise in AM and DfAM may play a deterministic role in the outcomes. That is, the designers' lack of expertise inhibits their ability to acknowledge "good" designs but not "bad" designs. For the latter, the DfAM worksheet may not be sensitive enough to elicit differences in the designs for those that experts scored highly. However, this may also tie into the designers' lack of expertise in DfAM and their interpretation of the DfAM concepts in the worksheet.

To promote expert-level DfAM reasoning, designers may require digital experiences that critically and comprehensively challenge their mental models of DfAM, beyond what was studied in this research. The added complexity of evaluating assemblies and multi-materials in design workflows may yield results in favor of added immersion [23,25]. Higher task complexity, such as evaluating manufacturability for a variety of print orientations and print parameters, may also identify more significant effects of immersion on DfAM evaluations. This research investigated design evaluation circumstances where such complexities were not present, likely influencing the observed findings or lack thereof. The limited scope may have also limited the observation of nuanced effects of immersion on DfAM outcomes that may be more apparent otherwise.

> *How do the differences in immersion between CAE, VR, and REAL modalities affect the cognitive load experienced when evaluating designs of varying manufacturability?*

The goal of RQ 2 was to understand how differences in immersion between modalities affect the cognitive load experienced from the DfAM evaluations. It was hypothesized that the VR and REAL evaluations would yield a lower cognitive load than the CAE evaluations. No significant differences between VR and REAL were expected. Regarding the primary effects of immersion, the results in Section 5.3 failed to reject the null hypothesis for each dimension. Specifically, the study could not identify significant differences in Perceptual, Response, Spatial, Visual, and Manual cognitive load between the conditions. These findings are interesting because this implies that the modality does not influence the effort experienced by designers processing information for DfAM evaluations. In other words, designers may find immersive and non-immersive mediums equally demanding (or comfortable) to evaluate designs for manufacturability by AM. Of particular interest here are the observations with the VR and REAL conditions. Note that participants in VR were generally exposed to a new environment, while those in CAE and REAL worked in familiar environments. Despite this, the results in Section 5.3 show that the cognitive load experienced by participants in VR was not significantly different from those in CAE and REAL. With the aid of a brief tutorial phase, this means that the novelty of the VR environment did not adversely influence the cognitive load experienced by the participants. This is interesting because it implies that DfAM evaluations in VR are as cognitively intuitive as those in CAE and REAL, even for novice or first-time users. Regarding the REAL condition, the data further suggests that manufacturability evaluations for AM may not merit the transition from digital to physical artifacts. In the broader scope, this presents interesting implications for how organizations create design workflows for AM.

> *How do the differences in engagement between CAE, VR, and REAL modalities explain the observed trends in DfAM evaluation outcomes and cognitive load?*

The goal of RQ 3 was to understand how differences in immersion between modalities affect engagement with the designs during the DfAM evaluations. It was hypothesized that analyzing how designers engage with the designs will explain the trends observed in RQ 1 and RQ 2. Specifically, by observing how immersion alters the perception of 3D artifacts and other visual information [73,74] and influences the interactions involved. The results in Section 5.4, however, present interesting findings that offer key context to the observations in Section 5.2 and Section 5.3 and their implications. Findings suggest that the modality for DfAM evaluations does not influence the passive engagement with the designs. This means that the designs retained their role as passive artifacts for communication, learning, and decision-making. In other words, designers across the conditions visualized the designs identically to extract information and make decisions. The findings also suggest that the modality strongly influences active engagement with the designs, specifically, when comparing the immersive and non-immersive modalities. Results showed that participants in CAE generally manipulated the designs more than those in VR and REAL. This means that designers in CAE were more likely to interact with the designs to extract information and make decisions.

These findings have strong implications for how a designer's engagement with designs may influence their DfAM evaluation outcomes and cognitive load. The collective findings from Sections 5.2 and 5.4 imply that high active engagement in non-immersive modalities is required to yield comparable DfAM outcomes to immersive modalities. Active engagement in non-immersive modalities may similarly curb experiencing higher cognitive load than in immersive modalities. These implications suggest immersive DfAM evaluations may not be constrained by the amount of active engagement with the designs, while non-immersive evaluations may be. It should be noted, however, that the comparison between VR and REAL showed an emerging trend (i.e., not a statistically significant trend) for differences in *Active* engagement. That is, participants may manipulate REAL designs more than those in VR, but not as much as those in CAE. A larger sample size could better identify trends between VR and REAL as well; however, it is unlikely to change the overall trend of the findings. Instead, an explanation for this emerging trend in *Active* engagement may be tied to the emerging trend for the effects of *Condition* on *Confidence*. Table 3a shows that the confidence reported by the participants seems to decrease as the condition changes from CAE to VR to REAL. A closer inspection of Fig. 9 shows that participants seemed more confident in their evaluations in CAE and VR than in REAL. This trend suggests that designers may be more confident evaluating digital designs over their physical counterparts. This means that the potentially higher manipulation of the REAL artifacts may be attributed to a lack of confidence in their evaluations.

An explanation for the emerging trend in *Confidence* and its potential relationship to *Active* engagement may be revealed by comparing the VR and REAL think-aloud recordings. First, the recordings showed participants in the REAL condition expressing more uncertainty in their evaluations. Since they were not given any information on the fabrication process and were novices in AM and DfAM, their mental models for manufacturability by AM may have been challenged by witnessing the fabricated artifacts. This phenomenon may have influenced the participants to manipulate the physical artifacts more than expected, perhaps to ascertain the fabrication process, contributing to the higher *Active* engagement in the REAL condition. Second, a further examination of the recordings for participants in the VR condition showed participants moving around the 3D model more than manually manipulating it. Specifically, the authors observed participants picking up the 3D model, manipulating it, and then suspending it in free space. Interestingly, they would then switch between moving around the model and moving the model around, with the former being more frequent than anticipated. While still recorded as *Active* engagements, they were generally observed to be brief, further contributing to the emerging trend for *Active* engagement. This suggests that the participants may have been more comfortable moving around the 3D model in VR than manipulating it. Whereas those in the REAL condition may have been more comfortable manipulating the physical artifacts than moving around them. Controlling for the uncertainty from DfAM evaluations of physical artifacts may present an interesting implication for how designers actively engage with the designs in immersive modalities. Specifically, digital immersion may induce more *non-manipulative* active engagement whereas physical immersion may induce more *manipulative* active engagement with designs.

Of additional interest in the observed findings is the lack of influence of higher active engagement in CAE on cognitive load. It makes sense that designers must frequently manipulate the 3D models to better evaluate them in non-immersive modalities. Given that cognitive load also does not vary with immersion (Section 5.3), habitual familiarity with the modality may play a key role in the effort exerted by the designer. That said, Table 4 does show an emerging trend for the effect of *Condition* on *Visual* cognitive load. A closer inspection of Fig. 10 shows a decrease in *Visual* cognitive load progressing linearly by the level of immersion. Specifically, the *Visual* cognitive load reported by the participants decreased as the condition changed from CAE to VR to REAL. Note that *Visual* processing cognitive load requires using attentional resources to process and interpret the meaning of visual information gained through sight. For example, seeing a sign on the road and comprehending what that means is an example where visual processing is used. In the DfAM evaluations, designers visually process a design's features and the information on the DfAM worksheet. This is a prerequisite for decision-making, which would trigger *Perceptual* processing cognitive load. However, the *Perceptual* cognitive load did not vary with immersion as shown in Table 4. This could mean that the level of immersion may not affect the decision-making but may affect the processing of visual information to make those decisions.

7. CONCLUSION

The presented work studied the design of VR experiences for DfAM applications. The goal was to understand how differences in immersion between modalities affect 1. the outcomes from the DfAM evaluations, 2. the cognitive load experienced from the evaluations, and 3. the engagement with the designs during the evaluations. A mixed-methods study was designed to extract quantitative and qualitative insights from the experiences of designers to inform this understanding. Participants evaluated multiple designs for manufacturability by ME in immersive and non-immersive modalities. Results suggest that outcomes from the DfAM evaluations do not observably vary with the immersion level of the modality. However, the relationship between the immersion level and the outcomes was found to be dependent on the design being evaluated. Additionally, the cognitive load experienced from the evaluations does not vary with immersion; however, emerging trends were observed. Furthermore, the results suggest that a designer's passive engagement with designs does not vary with immersion, but their active engagement does.

These contributions have significant implications for how future designers are trained in DfAM to meet the

AM demands in the workforce. Specifically, the findings suggest that immersive and non-immersive mediums can be used to train designers in DfAM without affecting their evaluation outcomes and experienced cognitive load. This work also presents interesting implications for how organizations create design workflows for AM. Specifically, the findings suggest that design for AM processes may not have a strong requirement to transition to physical artifacts. Instead, their digital counterparts may suffice for manufacturability evaluations, regarding AM processes like ME. However, designers may need to actively engage with the designs in digitally non-immersive modalities to achieve similar outcomes as those in digitally immersive modalities.

While these are interesting implications for DfAM applications, these findings must be considered with certain limitations of this work. This research limited its scope toward manufacturability evaluation for ME. Material extrusion is a relatively more accessible and functionally less complex process than processes like powder bed fusion. Therefore, the findings from this work may not be generalizable to other AM processes. Future work must expand on these findings and explore learning and intuition development for multiple AM processes. Doing so will aid industries in improving their digital design processes by empowering their designers with insight into the range of AM solutions. Additionally, the DfAM exercise in this work was limited to only visually evaluating designs for AM in a fixed print orientation. Participants did not have the opportunity to manipulate the designs to explore alternative print orientations. They also did not assess the impact of their decisions on the manufacturability outcomes of the designs, such as print time and support usage. Future work must incorporate a design problem that encourages designers to explore alternative print orientations when evaluating designs for AM. Furthermore, this research limited its scope to designers who were novices in AM and DfAM. Such findings may not be generalizable to veterans in the AM industry. Future work must explore the effects of immersion on designers with varying levels of expertise in AM and DfAM. Insight from such work could inform the design of experiences that are tailored to the expertise and needs of the designer.

Similarly, this work studied designers with little to no experience with VR. The observed findings were documented with designers who were habitually familiar with CAE and REAL modalities. Future work must study the role of modality familiarity on the observed findings, as well as investigate the effects of immersion on designers with similar levels of familiarity across CAE, VR, and REAL modalities. Finally, this research limited the scope of the think-aloud analysis to active and passive engagement. However, the subjective nature of how designers think and reason during DfAM evaluations leaves room to incorporate more sophisticated methods of analysis. Consider the emerging trend for *Active* engagement. Note that the recordings for participants in the REAL condition were cropped to anonymize the participants. This means that the raters had a constrained view of the participants' engagement with the designs. Although *fidgeting* was controlled when coding for *active engagement*, limited visibility limits the accuracy of such control. Future work must incorporate more sophisticated methods of analysis to account for such limitations. This could include analyzing eye-tracking and electroencephalography (EEG) data to further examine the designers' mental models and cognitive processes. Future work must also conduct a broader and deeper assessment of a designer's evaluation and decision-making processes to inform the design of immersive DfAM experiences.

ACKNOWLEDGMENTS

This research was conducted with the support of the National Science Foundation under Grant No. 2021267. Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF. We would also like to thank Dr. Stephanie Cutler for her continued guidance in working with undergraduate students for this initiative.

NOMENCLATURE

AM: Additive Manufacturing DfAM: Design for Additive Manufacturing ME: Material Extrusion CAE: Computer-Aided Engineering VR: Virtual Reality REAL: Physical (real) world

SUPPLEMENTARY DATA

Data collected from participants in this study is available upon request.

8. REFERENCES

[1] Gibson, I., Gao, Z., and Campbell, I., 2004, "A Comparative Study of Virtual Prototyping and Physical Prototyping," IJMTM, **6**(6), p. 503. DOI 10.1504/ijmtm.2004.005931.

[2] Zorriassatine, F., Wykes, C., Parkin, R., and Gindy, N., 2003, "A Survey of Virtual Prototyping Techniques for Mechanical Product Development," Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, **217**(4), pp. 513–530. DOI 10.1243/095440503321628189.

[3] Ford, S., and Despeisse, M., 2016, "Additive Manufacturing and Sustainability: An Exploratory Study of the Advantages and Challenges," Journal of Cleaner Production, **137**, pp. 1573–1587. DOI 10.1016/j.jclepro.2016.04.150.

[4] Ford, S., and Minshall, T., 2019, "Invited Review Article: Where and How 3D Printing Is Used in Teaching and Education," Additive Manufacturing, **25**, pp. 131–150. DOI 10.1016/j.addma.2018.10.028.

[5] Alfaify, A., Saleh, M., Abdullah, F. M., and Al-Ahmari, A. M., 2020, "Design for Additive Manufacturing: A Systematic Review," Sustainability, **12**(19), p. 7936. DOI 10.3390/su12197936.

[6] Pei, E., Monzón, M., and Bernard, A., eds., 2019, *Additive Manufacturing – Developments in Training and Education*, Springer International Publishing, Cham. DOI 10.1007/978-3-319-76084-1.

[7] Prabhu, R., Bracken, J., Armstrong, C. B., Jablokow, K., Simpson, T. W., and Meisel, N. A., 2020, "Additive Creativity: Investigating the Use of Design for Additive Manufacturing to Encourage Creativity in the Engineering Design Industry," International Journal of Design Creativity and Innovation, **8**(4), pp. 198–222. DOI 10.1080/21650349.2020.1813633.

[8] Prabhu, R., Simpson, T. W., Miller, S. R., Cutler, S. L., and Meisel, N. A., 2021, "Teaching Designing for Additive Manufacturing: Formulating Educational Interventions That Encourage Design Creativity," 3D Printing and Additive Manufacturing. DOI 10.1089/3dp.2021.0087.

[9] Simpson, T. W., Williams, C. B., and Hripko, M., 2017, "Preparing Industry for Additive Manufacturing and Its Applications: Summary & Recommendations from a National Science Foundation Workshop," Additive Manufacturing, **13**, pp. 166–178. DOI 10.1016/j.addma.2016.08.002.

[10] Meisel, N., and Williams, C., 2015, "An Investigation of Key Design for Additive Manufacturing Constraints in Multimaterial Three-Dimensional Printing," Journal of Mechanical Design, **137**(11), p. 111406. DOI 10.1115/1.4030991.

[11] Thompson, M. K., Moroni, G., Vaneker, T., Fadel, G., Campbell, R. I., Gibson, I., Bernard, A., Schulz, J., Graf, P., Ahuja, B., and Martina, F., 2016, "Design for Additive Manufacturing: Trends, Opportunities, Considerations, and Constraints," CIRP Annals, **65**(2), pp. 737–760. DOI 10.1016/j.cirp.2016.05.004.

[12] Booth, J. W., Alperovich, J., Chawla, P., Ma, J., Reid, T. N., and Ramani, K., 2017, "The Design for Additive Manufacturing Worksheet," Journal of Mechanical Design, **139**(10), p. 100904. DOI 10.1115/1.4037251.

[13] Bracken, J., Pomorski, T., Armstrong, C., Prabhu, R., Simpson, T. W., Jablokow, K., Cleary, W., and Meisel, N. A., 2020, "Design for Metal Powder Bed Fusion: The Geometry for Additive Part Selection (GAPS) Worksheet," Additive Manufacturing, **35**, p. 101163. DOI 10.1016/j.addma.2020.101163.

[14] Noh, H., Park, K., Park, K., and Okudan Kremer, G. E., 2021, "Development of a Design for Additive Manufacturing Worksheet for Medical Casts," *Volume 5: 26th Design for Manufacturing and the Life Cycle Conference (DFMLC)*, American Society of Mechanical Engineers, Virtual, Online, p. V005T05A004. DOI 10.1115/detc2021-72103.

[15] Gross, J., Park, K., and Kremer, G. E. O., 2018, "Design for Additive Manufacturing Inspired by TRIZ," *Volume 4: 23rd Design for Manufacturing and the Life Cycle Conference; 12th International Conference on Micro- and Nanosystems*, American Society of Mechanical Engineers, Quebec City, Quebec, Canada, p. V004T05A004. DOI 10.1115/detc2018-85761.

[16] Shi, Y., Zhang, Y., Baek, S., De Backer, W., and Harik, R., 2018, "Manufacturability Analysis for Additive Manufacturing Using a Novel Feature Recognition Technique," Computer-Aided Design and Applications, **15**(6), pp. 941–952. DOI 10.1080/16864360.2018.1462574.

[17] Budinoff, H. D., and McMains, S., 2021, "Will It Print: A Manufacturability Toolbox for 3D Printing," Int J Interact Des Manuf, **15**(4), pp. 613–630. DOI 10.1007/s12008-021-00786-w.

[18] Zhang, Y., Harik, R., Fadel, G., and Bernard, A., 2019, "A statistical method for build orientation determination in additive manufacturing," Rapid Prototyping Journal, **25**(1), pp. 187–207. DOI 10.1108/RPJ-04-2018-0102.

[19] Cayley, A., Mathur, J., and Meisel, N. A., 2023, "Creation and Assessment of a Novel Design Evaluation Tool for Additive Manufacturing," Journal of Mechanical Design, **146**(012001). DOI 10.1115/1.4063566.

[20] Blösch-Paidosh, A., and Shea, K., 2019, "Design Heuristics for Additive Manufacturing Validated Through a User Study1," Journal of Mechanical Design, **141**(4), p. 041101. DOI 10.1115/1.4041051.

[21] Lauff, C. A., Perez, K. B., Camburn, B. A., and Wood, K. L., 2019, "Design Principle Cards: Toolset to Support Innovations With Additive Manufacturing," *Volume 4: 24th Design for Manufacturing and the Life Cycle Conference; 13th International Conference on Micro- and Nanosystems*, American Society of Mechanical Engineers, Anaheim, California, USA, p. V004T05A005. DOI 10.1115/detc2019-97231.

[22] Berni, A., and Borgianni, Y., 2020, "Applications of Virtual Reality in Engineering and Product Design: Why, What, How, When and Where," Electronics, **9**(7), p. 1064. DOI 10.3390/electronics9071064.

[23] Horvat, N., Škec, S., Martinec, T., Lukačević, F., and Perišić, M. M., 2019, "Comparing Virtual Reality and Desktop Interface for Reviewing 3D CAD Models," Proc. Int. Conf. Eng. Des., **1**(1), pp. 1923–1932. DOI 10.1017/dsi.2019.198.

[24] Feeman, S. M., Wright, L. B., and Salmon, J. L., 2018, "Exploration and Evaluation of CAD Modeling in Virtual Reality," Computer-Aided Design and Applications, **15**(6), pp. 892–904. DOI 10.1080/16864360.2018.1462570.

[25] Wolfartsberger, J., 2019, "Analyzing the Potential of Virtual Reality for Engineering Design Review," Automation in Construction, **104**, pp. 27–37. DOI 10.1016/j.autcon.2019.03.018.

[26] Guo, Z., Zhou, D., Chen, J., Geng, J., Lv, C., and Zeng, S., 2018, "Using Virtual Reality to Support the Product's Maintainability Design: Immersive Maintainability Verification and Evaluation System," Computers in Industry, **101**, pp. 41–50. DOI 10.1016/j.compind.2018.06.007.

[27] Fillingim, K. B., Nwaeri, R. O., Paredis, C. J. J., Rosen, D., and Fu, K., 2020, "Examining the Effect of Design for Additive Manufacturing Rule Presentation on Part Redesign Quality," Journal of Engineering Design, **31**(8-9), pp. 427–460. DOI 10.1080/09544828.2020.1789569.

[28] Mathur, J., Miller, S. R., Simpson, T. W., and Meisel, N. A., 2022, "Identifying the Effects of Immersion on Design for Additive Manufacturing Evaluation of Designs of Varying Manufacturability," *Volume 5: 27th Design for Manufacturing and the Life Cycle Conference (DFMLC)*, American Society of Mechanical Engineers, St. Louis, Missouri, USA, p. V005T05A003. DOI 10.1115/detc2022-90063.

[29] Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I., 2020, "A Systematic Review of Immersive Virtual Reality Applications for Higher Education: Design Elements, Lessons Learned, and Research Agenda," Computers & Education, **147**, p. 103778. DOI 10.1016/j.compedu.2019.103778.

[30] Jensen, L., and Konradsen, F., 2018, "A Review of the Use of Virtual Reality Head-Mounted Displays in Education and Training," Educ Inf Technol, **23**(4), pp. 1515–1529. DOI 10.1007/s10639-017-9676-0.

[31] Slater, M., and Wilbur, S., 1997, "A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments," Presence: Teleoperators & Virtual Environments, **6**(6), pp. 603–616. DOI 10.1162/pres.1997.6.6.603.

[32] Vergara, D., Extremera, J., Rubio, M. P., and Dávila, L. P., 2019, "Meaningful Learning Through Virtual Reality Learning Environments: A Case Study in Materials Engineering," Applied Sciences, **9**(21), p. 4625. DOI 10.3390/app9214625.

[33] Violante, M. G., Vezzetti, E., and Piazzolla, P., 2019, "How to Design a Virtual Reality Experience That Impacts the Consumer Engagement: The Case of the Virtual Supermarket," Int J Interact Des Manuf, **13**(1), pp. 243–262. DOI 10.1007/s12008-018-00528-5.

[34] Richir, S., Fuchs, P., Lourdeaux, D., Millet, D., Buche, C., and Querrec, R., 2015, "How to Design Compelling Virtual Reality or Augmented Reality Experience?" International Journal of Virtual Reality, **15**(1), pp. 35–47. DOI 10.20870/ijvr.2015.15.1.2867.

[35] Li, K., Hall, M., Bermell-Garcia, P., Alcock, J., Tiwari, A., and González-Franco, M., 2017, "Measuring the Learning Effectiveness of Serious Gaming for Training of Complex Manufacturing Tasks," Simulation and Gaming, **48**(6), pp. 770–790. DOI 10.1177/1046878117739929.

[36] Perini, S., Luglietti, R., Margoudi, M., Oliveira, M., and Taisch, M., 2018, "Learning and Motivational Effects of Digital Game-Based Learning (DGBL) for Manufacturing Education –The Life Cycle Assessment (LCA) Game," Computers in Industry, **102**, pp. 40–49. DOI 10.1016/j.compind.2018.08.005.

[37] Chang, C.-C., Warden, C. A., Liang, C., and Lin, G.-Y., 2018, "Effects of Digital Game-Based Learning on Achievement, Flow and Overall Cognitive Load," Australasian Journal of Educational Technology, **34**(4). DOI 10.14742/ajet.2961.

[38] Despeisse, M., 2018, “GAMES AND SIMULATIONS IN INDUSTRIAL ENGINEERING EDUCATION: A REVIEW OF THE COGNITIVE AND AFFECTIVE LEARNING OUTCOMES,” *2018 Winter Simulation Conference (WSC)*, pp. 4046–4057. DOI 10.1109/wsc.2018.8632285.

[39] Tseng, T.-L., Pan, R., Zheng, J., Awalt, C., Gonzalez, M., and Medina, F., 2011, “Digital Additive Manufacturing for Engineering Education: A Virtual Rapid Prototyping Simulator Approach,” *2011 ASEE Annual Conference & Exposition Proceedings*, ASEE Conferences, Vancouver, BC, pp. 22.501.1–22.501.15. DOI 10.18260/1-2--17782.

[40] Deshpande, A. A., and Huang, S. H., 2011, “Simulation Games in Engineering Education: A State-of-the-Art Review,” Comput. Appl. Eng. Educ., **19**(3), pp. 399–410. DOI 10.1002/cae.20323.

[41] Ouyang, S.-G., Wang, G., Yao, J.-Y., Zhu, G.-H.-W., Liu, Z.-Y., and Feng, C., 2018, “A Unity3D-Based Interactive Three-Dimensional Virtual Practice Platform for Chemical Engineering,” Comput Appl Eng Educ, **26**(1), pp. 91–100. DOI 10.1002/cae.21863.

[42] Vlachopoulos, D., and Makri, A., 2017, “The Effect of Games and Simulations on Higher Education: A Systematic Literature Review,” International Journal of Educational Technology in Higher Education, **14**(1), p. 22. DOI 10.1186/s41239-017-0062-1.

[43] Berg, L. P., and Vance, J. M., 2017, “Industry Use of Virtual Reality in Product Design and Manufacturing: A Survey,” Virtual Reality, **21**(1), pp. 1–17. DOI 10.1007/s10055-016-0293-9.

[44] Houzangbe, S., Masson, D. H., Fleury, S., Gómez Jáuregui, D. A., Legardeur, J., Richir, S., and Couture, N., 2022, “Is Virtual Reality the Solution? A Comparison Between 3D and 2D Creative Sketching Tools in the Early Design Process,” Frontiers in Virtual Reality, **3**, p. 958223. DOI 10.3389/frvir.2022.958223.

[45] Yang, E. K., and Lee, J. H., 2020, “Cognitive Impact of Virtual Reality Sketching on Designers’ Concept Generation,” Digital Creativity, **31**(2), pp. 82–97. DOI 10.1080/14626268.2020.1726964.

[46] Meyer, O. A., Omdahl, M. K., and Makransky, G., 2019, “Investigating the Effect of Pre-Training When Learning Through Immersive Virtual Reality and Video: A Media and Methods Experiment,” Computers & Education, **140**, p. 103603. DOI 10.1016/j.compedu.2019.103603.

[47] Lovreglio, R., Duan, X., Rahouti, A., Phipps, R., and Nilsson, D., 2021, “Comparing the Effectiveness of Fire Extinguisher Virtual Reality and Video Training,” Virtual Reality, **25**(1), pp. 133–145. DOI 10.1007/s10055-020-00447-5.

[48] Krokos, E., Plaisant, C., and Varshney, A., 2019, “Virtual Memory Palaces: Immersion Aids Recall,” Virtual Reality, **23**(1), pp. 1–15. DOI 10.1007/s10055-018-0346-3.

[49] Hamilton, D., McKechnie, J., Edgerton, E., and Wilson, C., 2021, “Immersive Virtual Reality as a Pedagogical Tool in Education: A Systematic Literature Review of Quantitative Learning Outcomes and Experimental Design,” J. Comput. Educ., **8**(1), pp. 1–32. DOI 10.1007/s40692-020-00169-2.

[50] Grassini, S., Laumann, K., and Rasmussen Skogstad, M., 2020, “The Use of Virtual Reality Alone Does Not Promote Training Performance (but Sense of Presence Does),” Front. Psychol., **11**, p. 1743. DOI 10.3389/fpsyg.2020.01743.

[51] Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., and Bailenson, J. N., 2018, “Immersive Virtual Reality Field Trips Facilitate Learning About Climate Change,” Front. Psychol., **9**, p. 2364. DOI 10.3389/fpsyg.2018.02364.

[52] Makransky, G., Andreasen, N. K., Baceviciute, S., and Mayer, R. E., 2021, “Immersive Virtual Reality Increases Liking but Not Learning with a Science Simulation and Generative Learning Strategies Promote Learning in Immersive Virtual Reality.” Journal of Educational Psychology, **113**(4), pp. 719–735. DOI 10.1037/edu0000473.

[53] Klippel, A., Zhao, J., Oprean, D., Wallgrün, J. O., Stubbs, C., La Femina, P., and Jackson, K. L., 2020, “The Value of Being There: Toward a Science of Immersive Virtual Field Trips,” Virtual Reality, **24**(4), pp. 753–770. DOI 10.1007/s10055-019-00418-5.

[54] Huang, W., and Roscoe, R. D., 2021, “Head-Mounted Display-Based Virtual Reality Systems in Engineering Education: A Review of Recent Research,” Computer Applications in Engineering Education, **29**(5), pp. 1420–1435. DOI 10.1002/cae.22393.

[55] Pletz, C., and Zinn, B., 2020, “Evaluation of an Immersive Virtual Learning Environment for Operator Training in Mechanical and Plant Engineering Using Video Analysis,” British Journal of Educational Technology, **51**(6), pp. 2159–2179. DOI 10.1111/bjet.13024.

[56] Soliman, M., Pesyridis, A., Dalaymani-Zad, D., Gronfula, M., and Kourmpetis, M., 2021, “The Application of Virtual Reality in Engineering Education,” Applied Sciences, **11**(6), p. 2879. DOI 10.3390/app11062879.

[57] Parong, J., and Mayer, R. E., 2018, “Learning Science in Immersive Virtual Reality,” Journal of Educational Psychology, **110**, pp. 785–797. DOI 10.1037/edu0000241.

[58] Starkey, E. M., McKay, A. S., Hunter, S. T., and Miller, S. R., 2018, “Piecing Together Product Dissection: How Dissection Conditions Impact Student Conceptual Understanding and Cognitive Load,” Journal of Mechanical Design, **140**(5), p. 052001. DOI 10.1115/1.4039384.

[59] Wismer, A., Reinerman-Jones, L., Teo, G., Willis, S., McCracken, K., and Hackett, M., 2018, “A Workload Comparison During Anatomical Training with a Physical or Virtual Model,” *Augmented Cognition: Users and Contexts*, D.D. Schmorrow, and C.M. Fidopiastis, eds., Springer International Publishing, Cham, pp. 240–252. DOI 10.1007/978-3-319-91467-1_20.

[60] Armougum, A., Orriols, E., Gaston-Bellegarde, A., Marle, C. J.-L., and Piolino, P., 2019, “Virtual Reality: A New Method to Investigate Cognitive Load During Navigation,” Journal of Environmental Psychology, **65**, p. 101338. DOI 10.1016/j.jenvp.2019.101338.

[61] Frederiksen, J. G., Sørensen, S. M. D., Konge, L., Svendsen, M. B. S., Nobel-Jørgensen, M., Bjerrum, F., and Andersen, S. A. W., 2020, “Cognitive Load and Performance in Immersive Virtual Reality Versus Conventional Virtual Reality Simulation Training of Laparoscopic Surgery: A Randomized Trial,” Surg Endosc, **34**(3), pp. 1244–1252. DOI 10.1007/s00464-019-06887-8.

[62] Buxton, W., 2007, *Sketching User Experiences: Getting the Design Right and the Right Design*, Elsevier/Morgan Kaufmann, Amsterdam Boston.

[63] Sweller, J., van Merriënboer, J. J. G., and Paas, F., 2019, “Cognitive Architecture and Instructional Design: 20 Years Later,” Educ Psychol Rev, **31**(2), pp. 261–292. DOI 10.1007/s10648-019-09465-5.

[64] Barnawal, P., Dorneich, M. C., Frank, M. C., and Peters, F., 2017, “Evaluation of Design Feedback Modality in Design for Manufacturability,” Journal of Mechanical Design, **139**(9), p. 094503. DOI 10.1115/1.4037109.

[65] Boschma, R., 2005, “Proximity and Innovation: A Critical Assessment,” Regional Studies, **39**(1), pp. 61–74. DOI 10.1080/0034340052000320887.

[66] Koskinen, K. U., and Vanharanta, H., 2002, “The Role of Tacit Knowledge in Innovation Processes of Small Technology Companies,” International Journal of Production Economics, **80**(1), pp. 57–64. DOI 10.1016/S0925-5273(02)00243-8.

[67] Gibson, I., Rosen, D., Stucker, B., and Khorasani, M., 2021, *Additive Manufacturing Technologies*, Springer International Publishing, Cham. DOI 10.1007/978-3-030-56127-7.

[68] Jiang, R., Kleer, R., and Piller, F. T., 2017, “Predicting the Future of Additive Manufacturing: A Delphi Study on Economic and Societal Implications of 3D Printing for 2030,” Technological Forecasting and Social Change, **117**, pp. 84–97. DOI 10.1016/j.techfore.2017.01.006.

[69] Caviggioli, F., and Ughetto, E., 2019, “A Bibliometric Analysis of the Research Dealing with the Impact of Additive Manufacturing on Industry, Business and Society,” International Journal of Production Economics, **208**, pp. 254–268. DOI 10.1016/j.ijpe.2018.11.022.

[70] Kumke, M., Watschke, H., Hartogh, P., Bavendiek, A.-K., and Vietor, T., 2018, “Methods and Tools for Identifying and Leveraging Additive Manufacturing Design Potentials,” Int J Interact Des Manuf, **12**(2), pp. 481–493. DOI 10.1007/s12008-017-0399-7.

[71] Abidi, M. H., Al-Ahmari, A., Ahmad, A., Ameen, W., and Alkhalefah, H., 2019, “Assessment of Virtual Reality-Based Manufacturing Assembly Training System,” Int J Adv Manuf Technol, **105**(9), pp. 3743–3759. DOI 10.1007/s00170-019-03801-3.

[72] Bharathi, A. K. B. G., and Tucker, C. S., 2015, “Investigating the Impact of Interactive Immersive Virtual Reality Environments in Enhancing Task Performance in Online Engineering Design Activities,” *Volume 3: 17th International Conference on Advanced Vehicle Technologies; 12th International Conference on Design Education; 8th Frontiers in Biomedical Devices*, American Society of Mechanical Engineers, Boston, Massachusetts, USA, p. V003T04A004. DOI 10.1115/detc2015-47388.

[73] Paes, D., Irizarry, J., and Pujoni, D., 2021, “An Evidence of Cognitive Benefits from Immersive Design Review: Comparing Three-Dimensional Perception and Presence Between Immersive and Non-Immersive Virtual Environments,” Automation in Construction, **130**, p. 103849. DOI 10.1016/j.autcon.2021.103849.

[74] Tabbers, H. K., Martens, R. L., and Merriënboer, J. J. G., 2004, “Multimedia Instructions and Cognitive Load Theory: Effects of Modality and Cueing,” British Journal of Educational Psychology, **74**(1), pp. 71–81. DOI 10.1348/000709904322848824.

[75] Ibrahim, R., and Pour Rahimian, F., 2010, “Comparison of CAD and Manual Sketching Tools for Teaching Architectural Design,” Automation in Construction, **19**(8), pp. 978–987. DOI 10.1016/j.autcon.2010.09.003.

[76] Pontonnier, C., Dumont, G., Samani, A., Madeleine, P., and Badawi, M., 2014, “Designing and Evaluating a Workstation in Real and Virtual Environment: Toward Virtual Reality Based Ergonomic Design Sessions,” J Multimodal User Interfaces, **8**(2), pp. 199–208. DOI 10.1007/s12193-013-0138-8.

[77] Mathur, J., 2022, “DfAM Worksheet” [Online]. Available: https://sites.psu.edu/madebydesign/files/2017/07/worksheet.pdf.

[78] Tsang, P. S., and Velazquez, V. L., 1996, “Diagnosticity and Multidimensional Subjective Workload Ratings,” Ergonomics, **39**(3), pp. 358–381. DOI 10.1080/00140139608964470.

[79] Mathur, J., 2023, “3D Models for the DfAM Evaluation Exercise” [Online]. Available: https://sites.psu.edu/madebydesign/files/2023/10/assets.zip.

[80] Bradley, J. V., 1958, “Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design,” Journal of the American Statistical Association, **53**(282), pp. 525–528. DOI 10.1080/01621459.1958.10501456.

[81] Girard, J. M., and C. Wright, A. G., 2018, “DARMA: Software for Dual Axis Rating and Media Annotation,” Behav Res, **50**(3), pp. 902–909. DOI 10.3758/s13428-017-0915-5.

[82] Lauff, C. A., Kotys-Schwartz, D., and Rentschler, M. E., 2018, “What Is a Prototype? What Are the Roles of Prototypes in Companies?” Journal of Mechanical Design, **140**(6). DOI 10.1115/1.4039340.

[83] Bucciarelli, L. L., 2002, “Between Thought and Object in Engineering Design,” Design Studies, **23**(3), pp. 219–231. DOI 10.1016/S0142-694X(01)00035-7.

[84] Vinck, D., 2011, “Taking Intermediary Objects and Equipping Work into Account in the Study of Engineering Practices,” Engineering Studies, **3**(1), pp. 25–44. DOI 10.1080/19378629.2010.547989.

[85] James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., and Goodale, M. A., 2002, "'Active' and 'Passive' Learning of Three-Dimensional Object Structure Within an Immersive Virtual Reality Environment," Behavior Research Methods, Instruments, & Computers, **34**(3), pp. 383–390. DOI 10.3758/BF03195466.

[86] Fogarty, J., McCormick, J., and El-Tawil, S., 2018, "Improving Student Understanding of Complex Spatial Arrangements with Virtual Reality," Journal of Professional Issues in Engineering Education and Practice, **144**(2), p. 04017013. DOI 10.1061/(ASCE)EI.1943-5541.0000349.

[87] Rubio, S., Diaz, E., Martin, J., and Puente, J. M., 2004, "Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods," Applied Psychology, **53**(1), pp. 61–86. DOI 10.1111/j.1464-0597.2004.00161.x.

[88] Peña, E. A., and Slate, E. H., 2006, "Global Validation of Linear Model Assumptions," Journal of the American Statistical Association, **101**(473), pp. 341–354. DOI 10.1198/016214505000000637.

[89] Loy, A., and Hofmann, H., 2014, "**HLMdiag** : A Suite of Diagnostics for Hierarchical Linear Models in *R*," J. Stat. Soft., **56**(5). DOI 10.18637/jss.v056.i05.

List of Figures

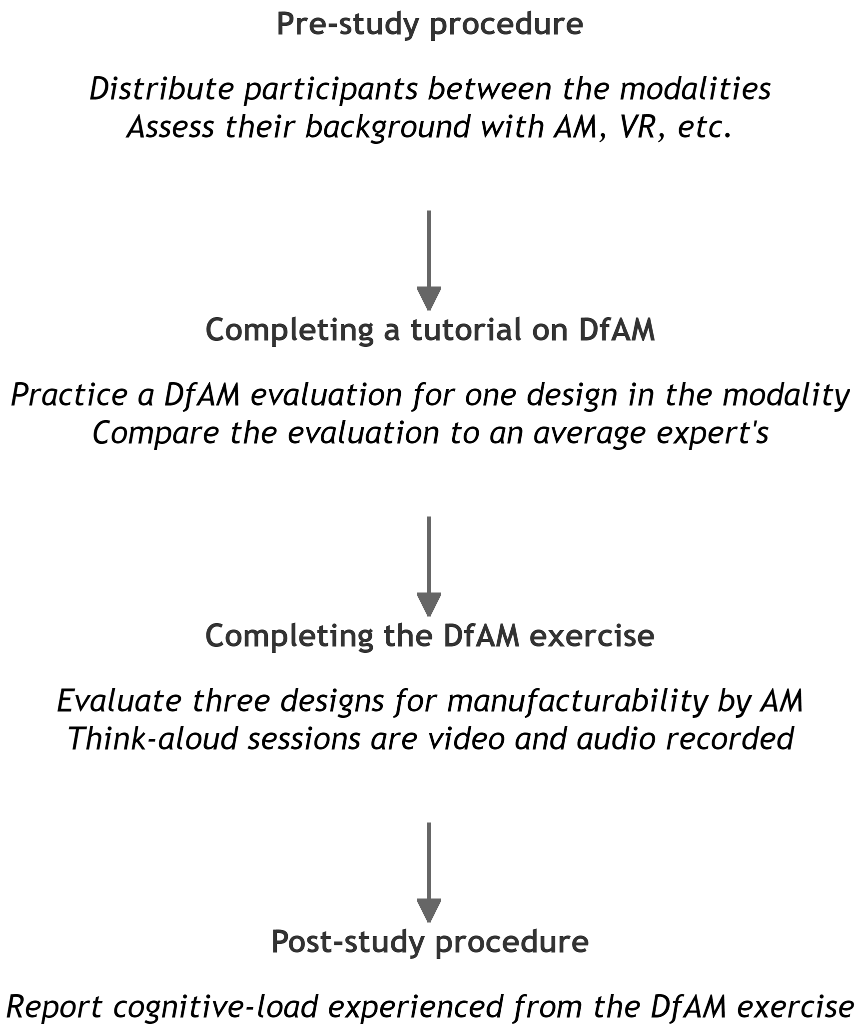
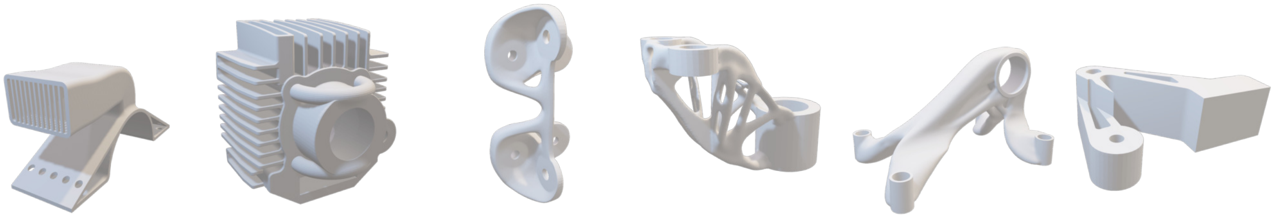
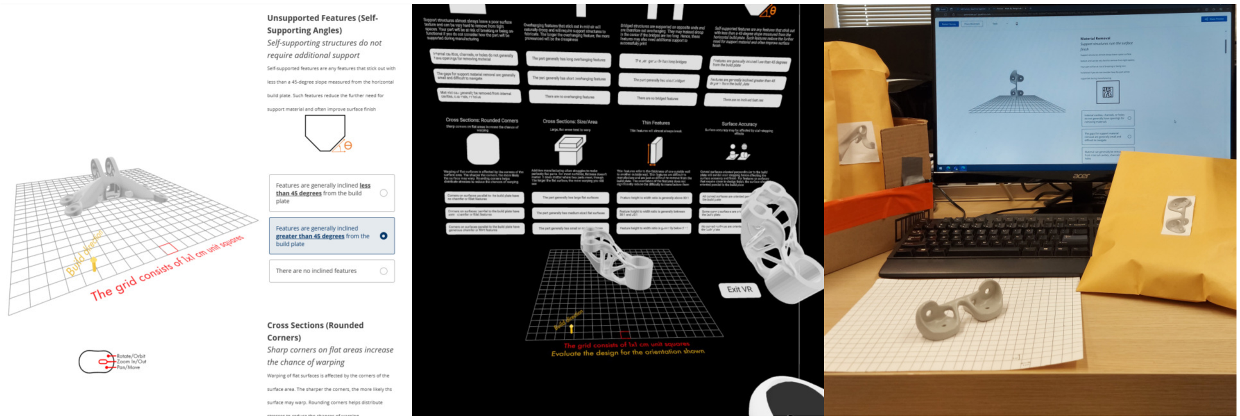


Fig. 1: Illustrating the steps for the designed mixed-methods experimentation



D1 (11.8) D2 (13.5) D3 (17.5) D4 (17.5) D5 (20) D6 (21.3)

Fig. 2: Displaying the designs selected for the DfAM evaluation exercise (with their expert-assigned DfAM scores)



Unsupported Features (Self-Supporting Angles)

Self-supporting structures do not require additional support

Self-supported features are any features that stick out with less than a 45-degree slope measured from the horizontal build plate. Such features reduce the further need for support material and often improve surface finish

Features are generally inclined **less than 45 degrees** from the build plate ○

Features are generally inclined **greater than 45 degrees** from the build plate ●

There are no inclined features ○

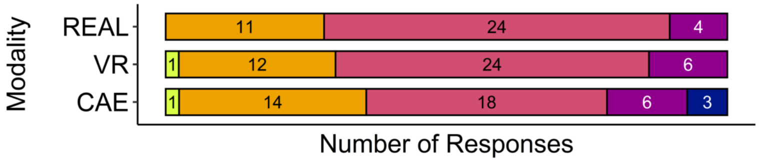
Cross Sections (Rounded Corners)

Sharp corners on flat areas increase the chance of warping

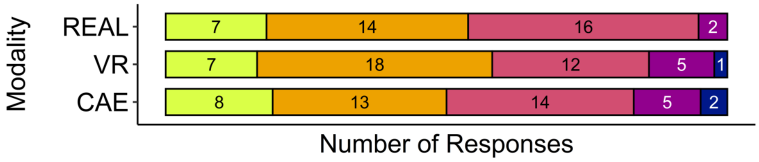
Warping of flat surfaces is affected by the corners of the surface area. The sharper the corners, the more likely ths surface may warp. Rounding corners helps distribute

(a) CAE setup (b) VR setup (c) REAL setup

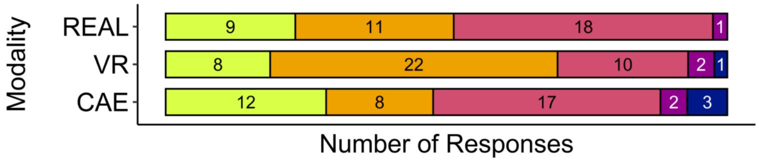
Fig. 3: Presenting the design of the DfAM evaluation environments for each condition. Each environment included a 3D artifact, tools for evaluating the design, and instructions for completing the exercise



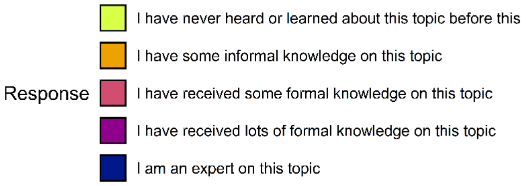
(a) Prior knowledge of AM



(b) Prior knowledge of ME



(c) Prior knowledge of DfME



(d) Common legend for responses

Fig. 4: Presenting the distribution of reported prior knowledge on AM, ME, and DfME between the conditions

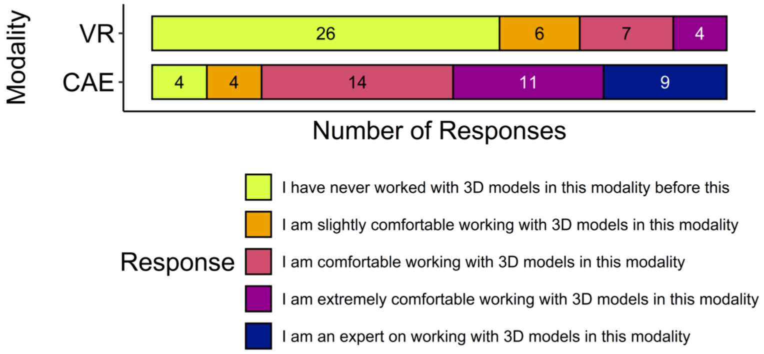


Fig. 5: Presenting the distribution of reported proficiency on working with CAE and VR modalities

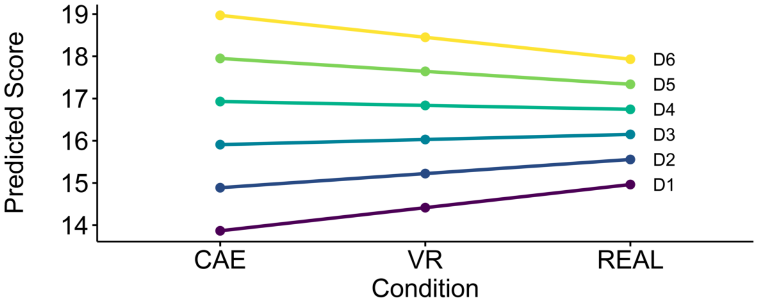


Fig. 6: Illustrating the interaction between the effects of condition and design on the DfAM scores

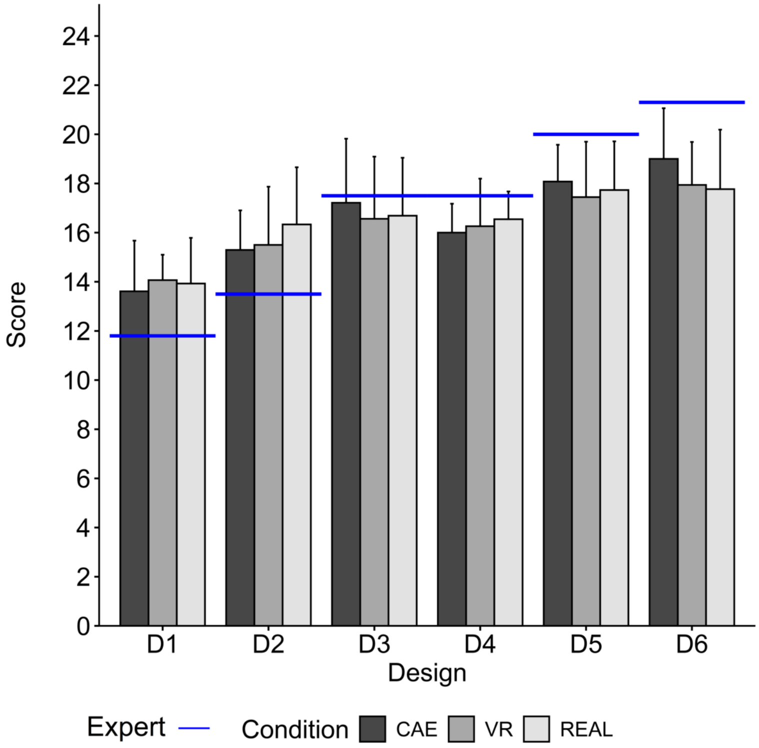


Fig. 7: Illustrating the DfAM scores assigned for each design across the conditions. A comparison to a condition-independent expert score for each design is also presented

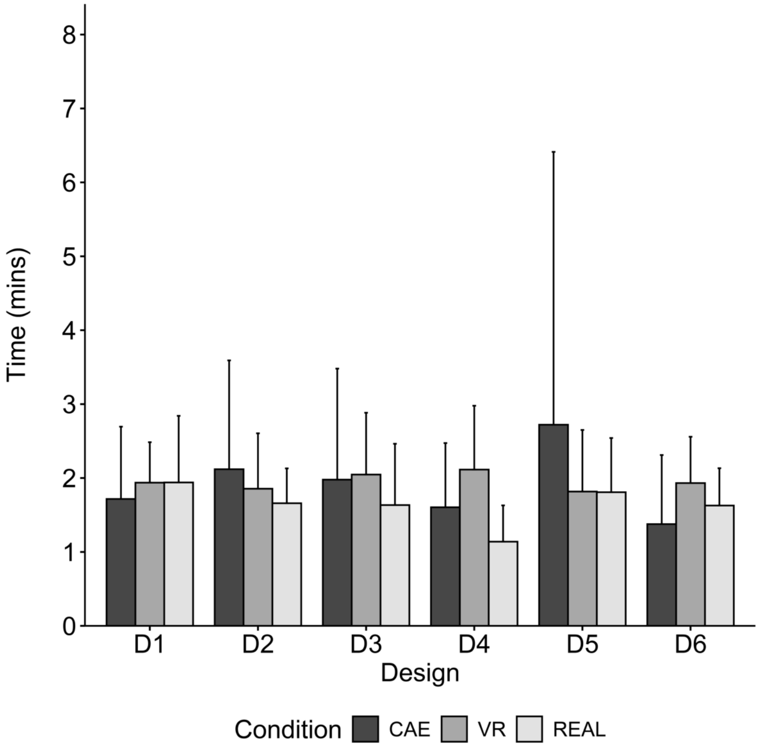


Fig. 8: Illustrating the time taken by participants to evaluate each design across the conditions

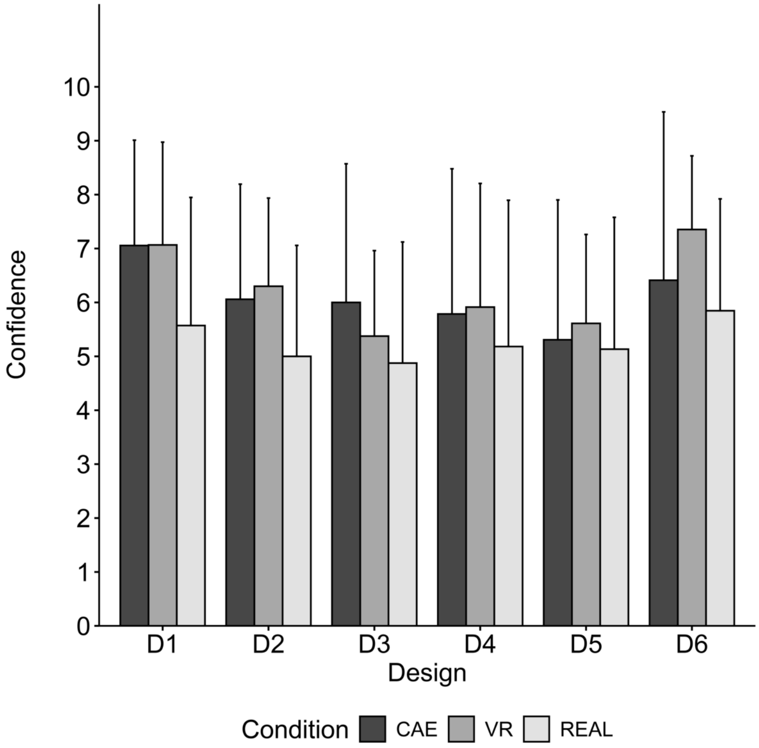


Fig. 9: Illustrating the confidence expressed on the DfAM evaluation for each design across the conditions

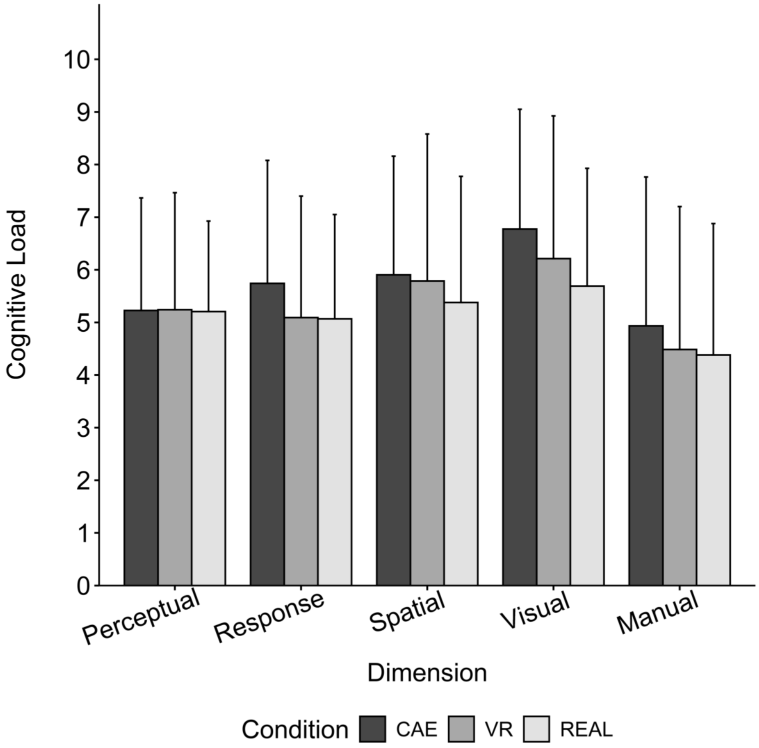


Fig. 10: Showing the distribution of reported cognitive load as affected by the three conditions

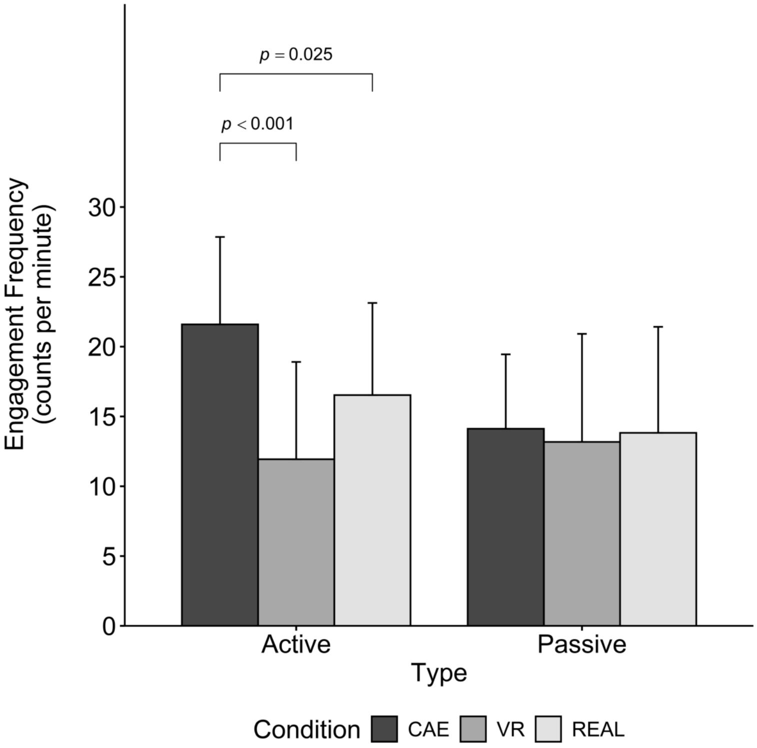


Fig. 11: Illustrating the observed frequency of engagement for each type across the conditions

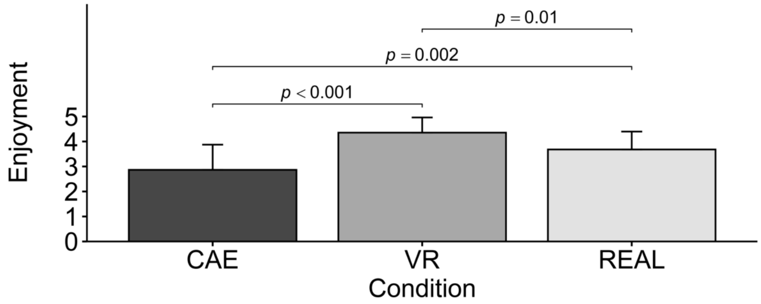


Fig. 12: Illustrating the reported *enjoyment* experienced from using each modality for DfAM evaluations

List of Tables

Table 1: Codebook used to analyze the video and audio recordings and identify emerging themes

Code	Description	Example
Referencing	General expressions or actions on recognizing shapes and features, picturing or imagining features and objects, or making estimations and assumptions to aid in evaluative decision-making	“This clearly has a lot of overhanging features” “I don’t think I see any bridged features” “I can see these overhangs over here needing a lot of support material, but it should be easy to remove” “I think these look twice as high as they are wide”
Interacting	General and intentional actions to manipulate the 3D model (or move around the 3D model) to evaluate it from different perspectives	Observable manipulation within or of the environment or the 3D model with a *clear intention to evaluate the design* (i.e., not simply fidgeting with the model or environment)

Table 2: Displaying the distribution of participants between the methods of study and the conditions

	CAE	VR	REAL
Qualitative	11	10	10
Quantitative	31	33	29

Table 3: Listing the different experimental variables and their statistical effect on the DfAM outcomes

(a) Effect of Condition

Outcome	Estimate	n.df	df	F.value	t.ratio	p.value
Score	0.03	1	91	0.01	0.08	0.936
Time	-0.24	1	91	1.28	-1.13	0.261
Confidence	-0.90	1	91	2.70	-1.64	0.104

(b) Effect of Design

Outcome	Estimate	n.df	df	F.value	t.ratio	p.value
Score	4.04	1	257	134.54	11.60	0.000
Time	0.01	1	254	0.00	0.03	0.974
Confidence	0.04	1	195	0.04	0.20	0.838

(c) Effect of the Condition and Design interaction

Outcome	Estimate	n.df	df	F.value	t.ratio	p.value
Score	-2.14	1	252	6.32	-2.51	0.013
Time	-0.20	1	249	0.18	-0.42	0.675
Confidence	0.50	1	194	1.05	1.02	0.307

Table 4: Listing the different cognitive load dimensions, indicating how they generally differ across the conditions

Dimension	Estimate	F(1, 91)	t.ratio	p.value
Perceptual	-0.02	0.00	-0.03	0.972
Response	-0.68	1.41	-1.19	0.238
Spatial	-0.52	0.66	-0.81	0.420
Visual	-1.09	3.02	-1.74	0.085
Manual	-0.56	0.66	-0.81	0.420

Table 5: Showing the Intraclass Correlation Coefficient (ICC) between the two raters from coding 6 recordings

PID	Condition	Evaluation	icc.coeff
1	REAL	Active	0.614
1	REAL	Passive	0.658
2	VR	Active	0.880
2	VR	Passive	0.788
3	CAE	Active	0.853
3	CAE	Passive	0.751
4	REAL	Active	0.788
4	REAL	Passive	0.676
5	VR	Active	0.778
5	VR	Passive	0.624
6	CAE	Active	0.730
6	CAE	Passive	0.734