D360: a Tool for Supporting Rapid, Iterative, and Collaborative Analysis of 360° Video

Meijer, Wo; Dingler, Tilman; Kortuem, Gerd

DOI
10.1145/3715336.3735793

Publication date
2025

Document Version
Final published version

Published in
DIS 2025 - Proceedings of the 2025 ACM Designing Interactive Systems Conference

D360: a Tool for Supporting Rapid, Iterative, and Collaborative Analysis of 360° Video

Wo Meijer
TU Delft
Faculty of Industrial Design Engineering
Delft, Netherlands
W.I.M.T.Meijer@tudelft.nl

Tilman Dingler
Delft University of Technology
Faculty of Industrial Design Engineering
Delft, Netherlands
t.dingler@tudelft.nl

Gerd Kortuem
Delft University of Technology
Faculty of Industrial Design Engineering
Delft, Netherlands
g.w.kortuem@tudelft.nl

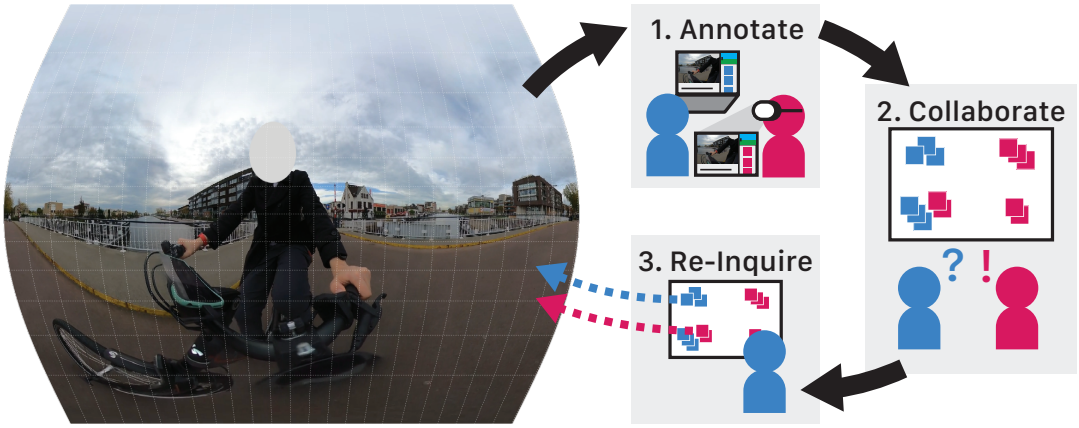


Figure 1: D360 is a tool that enables designers to utilize 360° video in their user research process through a web based viewing and annotation tool (the D360 Viewer). Annotations made with this tool are automatically inserted into Miro, enabling designers to collaboratively analyze the 360° video material. The annotations in Miro include a back-link to the D360 viewer, which lets designers re-inquire past insights as the scope and focus of the design process change.

Abstract

Designers can immerse themselves into the world of users by using 360° video leading to richer insights and better solutions. However, 360° video is challenging to share and incompatible with existing tools, preventing designers from effectively integrating it into their iterative and collaborative workflows. To address these challenges, we developed D360, a tool that enables designers to view, annotate, and collaboratively analyze 360° video. D360 features a web-based 360° video viewing and annotation tool, a database, and Miro integration to analyze 360° video using a familiar collaborative process. We evaluated D360 using walk-throughs with six professional designers that verified its utility and identified improvements to creating and presenting annotations. By providing both design directions for future 360° video tools for designers and our open source tool, we enable practitioners and researchers to leverage the rich interaction and visual context of 360° video for more impactful insights.

DIS '25, Funchal, Portugal

ACM ISBN 979-8-4007-1485-6/25/07
https://doi.org/10.1145/3715336.3735793

CCS Concepts

• **Human-centered computing** → *Walkthrough evaluations*; **Contextual design**; Collaborative interaction.

Keywords

360° Video, Video Design Ethnography, Contextual Inquiry

ACM Reference Format:
Wo Meijer, Tilman Dingler, and Gerd Kortuem. 2025. D360: a Tool for Supporting Rapid, Iterative, and Collaborative Analysis of 360° Video. In *Designing Interactive Systems Conference (DIS '25), July 05–09, 2025, Funchal, Portugal.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3715336.3735793

1 Introduction

Designers probe into the world of their users to inform the development of better products [33, 37, 46]. While designers use many different sources of information in their research, video specifically offers visually and temporally rich data, capturing complex interactions and behaviors over time [61, p.26]. However, video provides a limited window into the world of users - cameras have a limited Field of View (FoV) which can omit critical contextual information [49].

A technological solution for this shortcoming is 360° video [22, 32, 55]. 360° video provides a complete view of the visual context

around users [18] and enables more thorough analysis of how users act and interact in their environment [56]. For example, Neubauer et al. [34] explore how designers can use 360° video to understand the context of astronauts on the International Space Station to develop an empathic understanding of a context that is literally a different world than that of the designers. Despite its advantages, 360° video introduces new challenges. Conventional tools for video analysis lack support for spherical video, making it difficult to share and analyze 360° content [32]. Furthermore, existing 360° video tools are aimed at "flattening" 360° videos by selecting specific viewpoints, which does not support iterative exploration of 360° videos and encourages users to discard the 360° nature of the video for the convenience of conventional video.

To overcome these challenges, we developed the D360 system. First we used existing literature and first-hand experience to synthesize six design requirements that address challenges 360° video introduces during the stages of viewing, sharing, and re-inquiring that are crucial to Video Design Ethnography. D360 provides a web-based 360° video viewer where users can annotate 360° video. These annotations are then stored in a database along with information about the 360° video and users. Finally, to support collaborative analysis, D360 integrates with widely used online white-boarding tools such as Miro[1].

We evaluated the utility of the D360 system with an example scenario and walk-through sessions with 6 professional designers. The results demonstrate that the D360 system effectively facilitates the analysis of 360° video by giving designers a low-threshold interface for viewing and annotating 360° video, bringing those insights into a familiar interface (Miro), and providing a simple method to re-engage with the 360° video. Our analysis also indicates a number of additional design considerations for 360° video analysis tools for designers that will be integrated into a future version of D360. Furthermore, we discuss how tools like D360 can support the use of 360° video outside of design in use cases such as education which also focus on iterative engagement with material. However, our evaluations only explored artificial uses of D360 - in order to facilitate the study of D360 as used (or not) by designers in practice, as well as to support the extension or use of D360 for other purposes, we have open sourced D360 at [redacted for review].

By describing and evaluating the design considerations behind D360 as well as giving access of the tool to anyone, we aim to support the exploration and use of 360° video by designers, educators, laypeople, and researchers in order to better understand the actions, reactions, and contexts of users, professionals, and the subjects of ethnographic study.

2 Background

2.1 Design Ethnography

Nova [37, p.34] state that designers work with "the assumption [...] that documenting people's practices and products used in their natural habitat could be helpful for design". This assumption is the basis for approaches such as Contextual Design [5] – understanding the context of a user and use is an important factor in meaningful design decisions – or design probes [29] — inviting self-documentation

[1]https://miro.com/

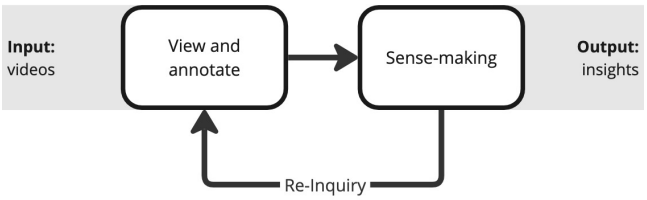


Figure 2: A generalized model for Video Design Ethnography adapted from Meijer et al. [32] that shows the iterative nature of viewing and annotating videos, collaboratively analyzing these annotations, and re-engaging with the material to challenge assumptions in the design process.

of users and their context to understand what elements of a design are important to them. This process of documenting people's practices is generalized as "design ethnography" [33, 37, 46], and centers around the goals of developing **insights** into the user, their context, and their behavior [11, 37, 45], **inspiration** for design concepts [14, 15], and **empathy** for (or understanding the internal state of) the user [17, 21, 53].

Importantly, the process of design ethnography is "abductive, constructivist, and reflexive" [33, p.21] meaning that the purpose and outcomes of analysis co-evolve with the interpretation of designers and the goals of the design process. Thus designers re-engage with user research multiple times with different levels of detail and intention. Because designers (often) work in teams [37, 61], this process happens both individually and between team members, with designers engaging in "...negotiation, collaboration, debate, conflict, and other social action" [61, p.34]. Therefore designers engage in "shared sense-making" [61, p.106], using bits of the research material that represent specific insights and are used, known as boundary objects [51], as a way to "negotiation of opinions in the team" [28]. These boundary objects often take the form of "stand-off annotations" [52] – annotations (usually text or drawings [31]) that exist outside of the media (e.g., videos, photos, etc.) they are annotating.

2.2 Video Design Ethnography (VDE)

There are many types of data that can be collected for Design Ethnography - interviews, observations, generative sessions, photos, videos, audio recordings, sketches, probes, etc. In this paper, we focus on the use of video as material for design ethnography, known as Video Design Ethnography (VDE). Video enables prolonged and unobtrusive observation of complex interactions [61, p.19], thereby enabling designers to develop insights into situations that would otherwise be too difficult, dangerous, or time consuming to observe. Beyond convenience, video also captures the temporal dynamics of interactions and contexts [61, p.26] which gives designers a window to understand not only a specific moment of a user's experience but how that moment came about and how those moments change over time. Video enables designers to continually re-engage with this rich and dynamic data, supporting an iterative process of discovering insights, negotiating the meaning of these insights, and then re-inquiring in the video to see if the new understanding of the insights matches the context [7]. Finally, designers use the visual

richness of individual frames of video to embody and share their insights [28]. These frames are then used to represent a shared understanding, often using online white-boarding tools that support remote and asynchronous collaboration [27]. This combination of visual richness, share-ability, and unobtrusiveness means video is well suited for design ethnography.

2.3 360° Video for Design Ethnography

While video provides rich and temporal information, it only provides a limited field of view into the world of users. This crucial limitation means that interactions and events happen "off camera", preventing designers from understanding all the actions and interactions in a context, for example not being able to understand where the users is looking during complex actions [49].

Fortunately, 360° action cameras[2] enable casual users of cameras to easily capture the full visual context around the camera [18]. Beyond simply capturing the context, the immersive nature of 360° video allows viewers to take the perspective of different actors in a scene, for example the view of a conductor or of an orchestra [56], and actively explore the video by moving their viewpoint, leading to greater immersion and empathy [2–4]. This additional context and immersion has been used to study a diverse set of contexts such as the International Space Station [34], firefighter training [47], dyeing practices [39], and diving [57, 58]. Thus demonstrating that 360° video not only eliminates the issue of framing the camera correctly, but it also gives designers the ability to engage and immerse themselves in contexts that they are totally unfamiliar with [32], enabling them to engage in what Westmoreland [58] calls "spatially aware analysis" which provides a richer understanding of the context and events within it.

However, Meijer et al. [32] discusses how the tools and techniques used for 360° video *ethnography* might not transfer readily to the use of 360° video by designers. This presents itself in two major ways, unlike conventional ethnography, designers collaboratively view [55] and analyze 360° videos, and they rapidly re-engage with the original material during discussions [7]. Meijer et al. [32] identified the need for a easily accessible, multi-device (i.e., VR headset, computer, and phone), tool that enables designers to annotate and view 360° video iteratively. This is distinct from the few tools developed for conventional ethnography using 360° video, such as AVA360VR [30] or CAVA360VR [41] used by Vatanen et al. [56], these focus on largely on VR headsets, which is how users both view and analyze 360° video – however, analysis in *design* ethnography is crucially an in-person collaborative discussion [61] that is hampered by participants needing to engage with technology [9].

2.4 Challenges Introduced by 360° Video

To dive deeper into the challenges 360° video introduces for design ethnography, it is important to understand that its spherical nature makes 360° content more difficult to view and share using the same techniques as conventional (flat) video [18, 32]. One approach to this challenge is to *flatten* 360° video by using any number of map projections [3]. However, this results in distorted images that are difficult for to understand [59] - especially the spatial relations between different actors in the video [20]. Therefore, many 360° video players show a conventional subset of the 360° video and let the viewer change that perspective, eliminating the impact of distortion.

However, this approach means that the viewer is responsible for controlling the view, which adds a new challenge of sharing or combining control when two or more people view collaboratively [18, 23] as well as introducing a fear of missing out on certain moments or areas of the video [1, 60]. The addition of a specific subset of the 360° video also adds an extra element to keep track of when viewers want to re-visit a specific moment – requiring the user to not only find the correct video and timestamp, but also the correct perspective [32, 35].

Finally, in order to share insights from 360° video, designers either need to present the full frame (introducing distortion), crop the frame (removing the visual context), or use a digital tool that lets the viewer change the perspective (which requires designers to disengage from discussions during collaborative workshops, which is undesirable [9]).

3 Design Goals

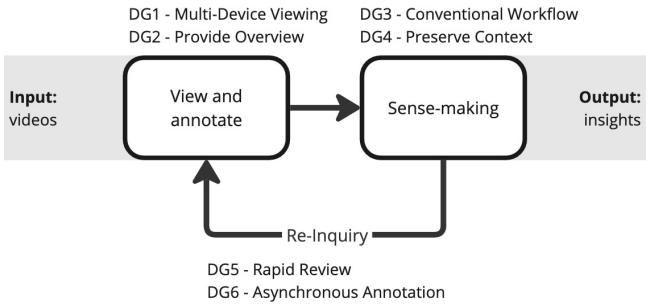


Figure 3: The six Design Goals used in the creation of D360 mapped onto a simplified version of the VDE process described in Section 2.2.

Based on the challenges that 360° video introduces to the process of VDE (Section 2.4), examples of 360° video tools from literature, as well as first hand experience using 360° video for design ethnography we synthesized six primary Design Goals (DGs) for tools that aim to support the use of 360° video by *designers*. While these could support other users of 360° video, designers especially consume 360° video collaboratively (unusual for conventional ethnography [55]), engage in collaborative sense-making [10, p.154], and often re-engage with the original (360°) video [7, 9] – the importance of these unique aspects are highlighted in Figure 3, which shows how these design goals link to the different stages of VDE.

- DG1 **Multi-Device Viewing**: Different methods of viewing 360° video offer different annotation experiences, such as the more immersive but cumbersome VR headset or the more analytical view of a laptop [32, 36]. Tools should support viewing 360° content using multiple devices to give users the freedom to select the appropriate tool.
- DG2 **Provide Overview**: Since users often only see a small area of a 360° video, which lead to stress about missing specific

[2]Such as the GoPro Max or Insta360 X4.
[3]For additional examples see: https://en.wikipedia.org/wiki/Map_projection

moments [32, 60]. Tools should enable users to maintain an overview of the full visual context to reduce this stress and increase contextual understanding.

DG3 **Conventional Workflow**: Designers have a workflow that they use in VDE to aid in engaging with collaborative analysis. Tools for 360° video analysis should integrate with workflows designers are familiar with [40] to simplify the process of adopting 360° video [32].

DG4 **Preserve Context**: One of the challenges with sharing insights from 360° video is the loss of the entire visual context when converting to flat screenshots [32]. Tools should represent annotations in a way that preserves both the visual context and orientation within the 360° view [23, 36].

DG5 **Rapid Review**: Finding the moment of a 360° video captured by a screenshot requires users to identify the correct video, time, and orientation, which is cumbersome [32, 36]. Tools should create "backlinks" that allow users to automatically return to the video, time, and orientation of annotations.

DG6 **Asynchronous Annotation**: Tools should support asynchronous collaboration (i.e., annotation by multiple designers) by attributing annotations to their creator, along with other meta-data [16, 54] (e.g., who used the tool, when, and what they created with the tool).

3.1 Existing 360° Video Tools

In the preceding sections we reference a few 360° video tools that help articulate and demonstrate the importance of some of the Design Guidelines, however these tools were made for other purposes than to support the unique process of 360° video design ethnography [32], and therefor it is logical that none of them fulfill all of the design goals. In this section we will discuss five 360° video tools that achieve *some* of the Design Goals – Vremiere [36], CollaVR [35], Tourgether360 [23], AVA360VR [30], and CACA360VR [41] – and discuss how none meet all the design goals set out for a tool to support 360° Video Design Ethnography.

The first two – Vremiere [36] and CollaVR [35] – are both aimed at collaborative editing of 360° video. Therefor they both implement the ability to view, annotate, and re-view 360° video. Specifically, Vremiere enables multi-device viewing (3, VR headset and computer), provides an overview and preserves context using the little planet projection (3 & 3), and enables rapid review with markers (3) – however, its focus on editing means it does not fit into the conventional workflow, with their evaluation finding a preference towards the participants own workflows over that of the tool. CollaVR is aimed at asynchronous annotation and viewing of 360° video (3), it enables collaborative viewing of 360° video *only* in VR headsets (thus not meeting 3), but does support thumbnails that preserve some of the context of an annotation (3). More importantly, both tools are not publicly available, which reduces their utility for designers more-so than the fact they not meet all the Design Goals described above.

Tourgether360 [23] is a tool that lets two participants collaboratively explore 360° videos, which indicates that it could be used to collaboratively analyze 360° video by designers. It allows the two viewers to see each other in the 360° video, as well as to indicate an interesting area of the 360° video. Additionally, it provides a "mini-map" that shows an overhead view of the path of the 360° video (3). However, while these features are great for a rich, real-time collaborative viewing experience, it fails to support the iterative process of annotating 360° videos, using those annotations elsewhere, and then returning to the 360° video if necessary. It also only supports VR headsets (not meeting 3), does not preserve context in annotations (not meeting 3), and does not integrate into the workflow of designers (not that it was designed to, but not meeting 3).

Finally, AVA360VR [30], and CACA360VR [41] are tools made by the BigSoftVideo group[4] for the purpose of analyzing 360° video in conventional ethnography (as demonstrated by Vatanen et al. [56]). Both tools are capable, they enable viewers to merge multiple materials (360° video, conventional video, multiple audio sources, transcripts) into a project, enable annotation by drawing on the 360° video, offer different 2D representations (3), and even enable viewers to record a presentation using the 360° video. However, this wide range of features comes at a cost – running the software requires a powerful computer, and more importantly, much of the work that designers would normally do in-person, can now be done only in a VR headset (not meeting 3 and 3). Thus, while AVA360VR and CAVA360VR are powerful tools for the more rigorous and solitary [55] methods of ethnography, they do not support the workflow and specifically in-person collaboration [8, 9] that is found in design ethnography.

Therefore, while there are a lot of tools that support some of the Design Guidelines that we synthesized, we have demonstrated that none of the tools we discussed fulfill all of them – as well as why these tools are not easily adaptable to the workflow of designers.

4 The D360 System

Based on the opportunities for a 360° video analysis tool for video design ethnography, we created D360, which consists of three pieces of software: (1) the D360 viewer, (2) a database for storing annotation sessions, and (3) Miro [5] integration. Figure 4 shows how annotations taken using the D360 viewer are saved on the database and then placed in a timeline on a Miro board. These annotations provide "back links" to the D360 viewer at the moment they were taken- enabling designers to quickly re-view the full 360° video.

4.1 The D360 Viewer

The D360 Viewer features a landing page which provides the user with an ID and allows them to to join a team; giving them access to an overview of a set of the 360° videos. When the user selects a 360° video to annotate, they are presented with the main D360 Viewer interface (Figure 5a). This interface provides a large window to view and navigate the 360° video with a conventional FoV (5a-A). The entire 360° frame of the video is included in a mini-map (5a-B) in order to give an overview of the full 360° context (3 Provide Overview). The bottom bar (5a-C) allows the user to pause the video, create new annotations (using the "+" button) and features a timeline that shows the progress in the video and gives an overview of annotations. These annotations are also shown on a scrolling

[4]See https://github.com/BigSoftVideo.

[5]While there are other online white-boarding tools available Miro was used at the companies participants 1,2, 4, and 6 work for.

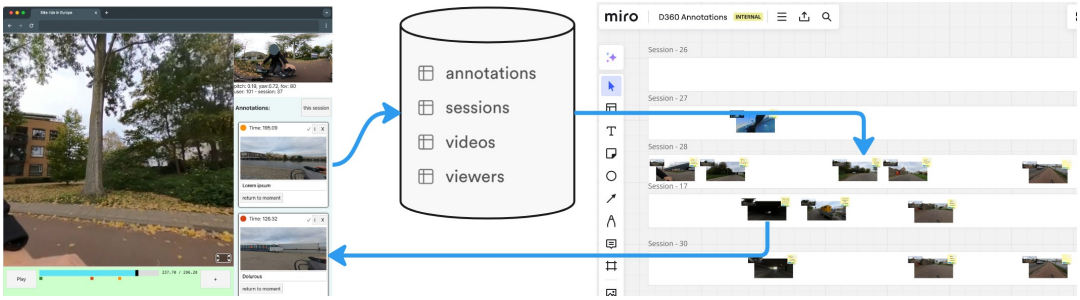


Figure 4: An overview of the D360 system consisting of the D360 viewer (left), the database (middle), and the integration into Miro (right).

sidebar (5a-D) which shows the annotation image and time and allows the user to add text to the annotation and return to that moment (time and perspective) of the 360° video. Finally, the D360 Viewer is built using A-Frame [6] which enables users to view and navigate 360° video on multiple devices (3 Multi-Device Viewing): computers (with mouse controls), phones (using touch and tilt controls), and VR headsets (using head tracking).

4.2 The Database

D360 uses a relational database to store information about the videos, users, annotations, and annotation sessions. Specifically, we used Supabase[7] which provides Javascript APIs, authentication, and file storage. The database stores the video ID, timestamp, orientation, and user ID of each annotation (3 Rapid Review). Additionally, the database creates an "annotation session" when a user starts annotating a video. This enables other viewers to view not just a singular annotations but contextualize how annotations from one user connect together (3 Asynchronous Annotation). Finally, the database enables annotations to act as back-links to the annotation session they were created in. This means that the annotation objects in Miro act as links to the full 360° video at the time and orientation of that annotation (3 Rapid Review).

4.3 Miro Integration

The data sent to the database can be integrated into a variety of tools (both for use by designers and for potential analysis of the annotating behavior of designers). Based on early input from the practitioners we engaged with, we opted to integrate the output of the D360 Viewer with Miro [8] - a popular online white-boarding tool that supports asynchronous and remote collaborative analysis. We created a small Miro plug-in that allows designers to quickly insert annotation objects (that include the full 360° visual context 3 Preserving Context) into Miro boards (3 Conventional Workflow). These annotation objects also contain the "back links" that connect the annotation back to the D360 viewer (to enable 3 Rapid Review).

The representation of annotations in Miro (Figure 5b) consists of three items grouped together as one object:

A The screenshot of the view when the annotation was created. With a link[9] that redirects the user to the D360 Viewer with the same video, timestamp, perspective, and annotation session as the associated annotation (enabling 3 Rapid Review).

B An overview of the full 360° frame to preserve the entire visual context of the annotation (enabling 3 Preserving Context).

C A sticky note with the text of the annotation that enables two way editing with the annotation object in the database. It is also tagged with the user that created the annotation and the name of the session the annotation was created in (to support 3 Asynchronous Annotation).

5 Evaluation

The goal of our evaluation is to understand the conceptual clarity, ease of use, and value to the intended audience of D360 (experienced designers using 360° video) [25]. Since VDE is a complex process with goals that vary with design team, context, and stage of the design process [61, p. 91], attempting to create standardized tasks in order to evaluate the real world usage of the D360 system would fall into the "usability trap" discussed by Olsen [38]. Therefore, we studied the usage of the D360 tool by designers with 360° video experience in a walk-through setting approaches, as recommended by Ledo et al. [25] to evaluate D360 for "utility, not usability". Additionally, we provide an example usage scenario (Appendix A) to help further illustrate the utility of D360 for design practitioners. We demonstrate both "…what the tool can do" [25] and the utility of D360 for its intended users. In other words, we address the following research questions:

RQ1 How do the Design Goals defined in Section 3 align with the experiences of designers?

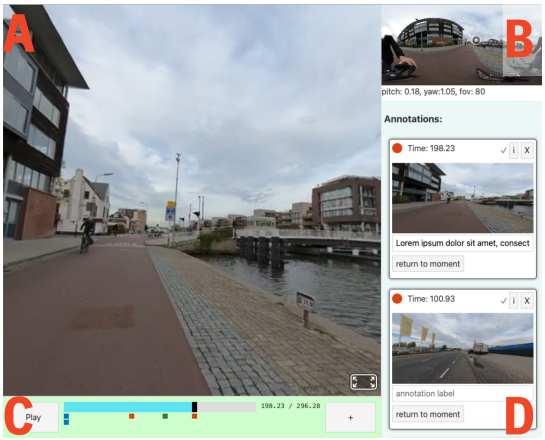
RQ2 How well are designers able to utilize the D360 tool to conduct VDE activities?

5.1 Participants

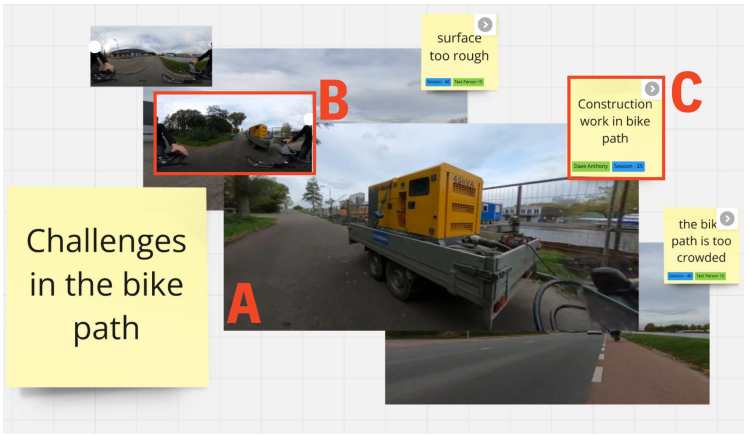
We conducted a walk-through evaluation [25, 44] with 6 designers who have experience working with 360° video for professional

[6] https://aframe.io/
[7] https://supabase.com/
[8] https://miro.com/

[9] Additionally, the image URL acts as a redirect, ensuring that the image - annotation link remains in the event of someone deleting the Miro link.



(a) The D360 Viewer web interface, showing (A) the 360° video view, (B) the mini-map, (C) the timeline and "+" button to create a new annotation, and (D) the scrollable list of annotations.



(b) An example of how D360 presents an annotation in Miro, composed of (A) the screenshot of the annotation, (B) an overview of the 360° frame, and (C) the text of the annotation.

Figure 5: (a) D360 Viewer interface, and (b) D360 annotation in Miro.

projects (Table 1). Participants were recruited through a combination of open calls on social media and snowball sampling. Participants were screened for experience with 360° video and engaged with 360° content[10] in their professional work - rejecting 2 potential candidates. The number of candidates was limited, mainly due to

	360° Experience	Design Context
P1	1 year	Robotics and Automation
P2	3 years	Urban Design
P3	3 years	Water Sports
P4	2 years	Urban Design
P5	4 years	ICU and Hospitals
P6	3 years	Robotics and Automation

Table 1: An overview of the participants, their experience with 360° photos and/or video, and the context of their professional work.

the novel nature of 360° video in ethnographic practice [32, 55]. We elected to use a smaller number of participants who represent the actual target audience of the tool, rather than a larger group of "stand-ins" to avoid the issues of transferability discussed by Ledo et al. [25].

5.2 Sessions

The sessions were conducted one-on-one and lasted approximately one hour[11], and participants were remunerated with coffee, tea, and snacks.

To demonstrate the abilities of D360 for asynchronous and iterative collaboration we selected an example design task of the lead researcher created a number of 360° videos of "identifying challenges and opportunities to integrate AI technology onto e-bikes". This design task was chosen as it aligned with participants experience with automation and urban environments. To avoid issues with privacy the lead researcher recorded a number of 360° videos of themselves cycling using a GoPro Max 360° action camera. These 360° videos were then annotated in order to pre-populate a Miro board that was used during the sessions. The sessions were structured in five phases:

(1) **Introduction:** the lead researcher introduces themselves to the participant, who then reads and signs an informed consent form.
(2) **Understanding the Participant's use of 360° Video:** The participant is asked to describe their workflow when using 360° video and discuss the benefits of 360° video as well as the challenges it presents.
(3) **Demonstrating the D360 viewer:** The D360 viewer interface is demonstrated for the participant, showing the features described in Figure 5a. The participant is asked to use the D360 viewer to annotate a brief segment of an example 360° video while being encouraged to think out loud.
(4) **Demonstrating the D360 Miro integration:** The participant is shown how the D360 tool integrates into Miro (Figure 5b), and is able to explore the Miro board with example annotations. The participant is encouraged to use the back link feature to see example annotations in the D360 viewer.
(5) **Critiquing the Utility of D360:** The participant is asked to reflect on the utility of D360, specifically elements of their process that are not addressed by the system or elements the system addresses that are extraneous.

5.2.1 Data and Analysis. All sessions were documented by screen recordings, an audio recorder, notes taken by the researcher, and notes or sketches made by the participant. This data was used

[10] Either 360° photos or 360° video.
[11] An unfortunately brief period since, limited by the availability of the participants. In many cases the discussion could have gone longer or the tool could have been used in multiple sessions.

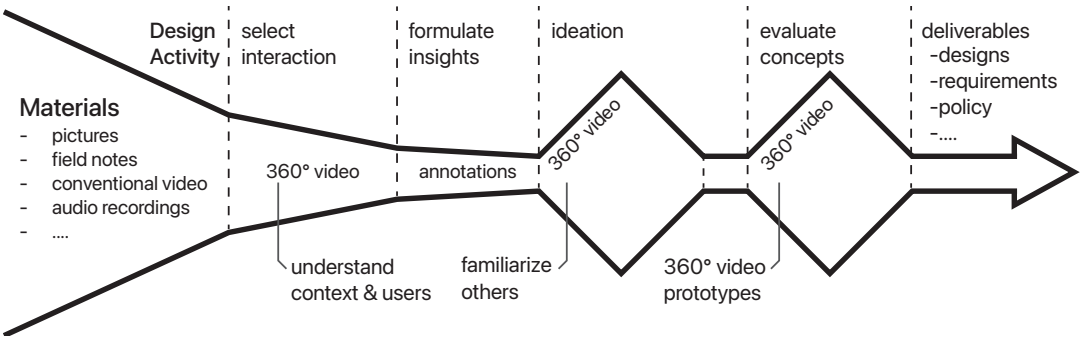


Figure 6: An amalgamation of sketches from P1, P5, and P6 describing how and when 360° video is used in their design process.

first analyzed to address the research questions by comparing the participants' descriptions of their process when using 360° video and summarizing their feedback related to how well D360 addresses the six design goals defined in Section 3. Later the data was used in an open, inductive process based on reflexive thematic analysis [6] in order to elicit novel design goals and unaddressed uses of 360° video in design. First the lead author familiarized themselves with the data and conducted initial coding of participant feedback and ideas. These initial codes then formed the basis of themes that aimed to combine key elements of the codes. These themes were then shared with the other authors who iteratively discussed and refined them.

6 Results

In this section we present the results of the walk-through sessions with professional designers. First, we describe how well the design goals formulated in Section 3 align with the experiences of designers (RQ1). Second, we describe three additional design goals elicited from the input of participants that were not described by previous work. Third, we describe participants' impression of the utility of the D360 system (RQ2). Finally, we define a number of additional features participants described for future versions of D360.

6.1 RQ1: alignment of design goals

All participants conducted analysis in a process similar to the reflexive VDE process described in Section 2.2, confirming the theoretical grounding of D360. Figure 6 shows an amalgamation of the process and how and when participants integrated 360° video.

For the specific design goals, participants described the importance and challenge of being able to find exact moments of 360° video 3 (P1, P3, P4, P5, P6), annotating asynchronously 3 (P1, P2, P3, P5), and frustration with integrating 360° content into their existing workflows 3 (P1, P2, P3, P4, P5, P6) without discarding the full visual context 3 (P1, P3, P4). Most participants limit their interaction with 360° video to laptops or phones due to limited availability of VR headsets (P1, P2, P4), but most (P1, P4, P5, P6) appreciated that different tools 3. Some participants expressed that they experienced feeling overwhelmed by the amount of visual information in 360° video 3 (P1, P3, P4), while the other participants did not mention it as a specific challenge with 360° video.

6.2 Additional Design Goals

Based on our analysis of participants' description of their workflow with 360° video in Phase 1 of the sessions as well as reflections during the critique of D360 in Phase 5, we extract three *additional* design goals (AGs) for tools that seek to aid designers' use of 360° video:

AG1 **Interaction Threshold:** One of the main challenges identified by participants is that the increased file size (P1, P3, P4, P6) of 360° videos and the need to use specific viewing software (P1, P2, P5, P6) presented significant friction when engaging with 360° video. P1 indicated that *"[they] can only watch 360° videos on [their] coworker's computer because he has the software"*. Similar logistical challenges include footage lost because SD cards were being used to transfer files from 360° cameras (P3, P6), incorrect camera settings resulting in 360° videos not being recorded (P4, P6), and file format issues (P1, P2). Largely this means that a lot of time is spent on the logistics of 360° videos both in analysis and when sharing the video with others during workshops (P6), and that remote collaboration with 360° videos is *"...limited by the software and devices that [team members] have access to"* (P4).

AG2 **Mixed Sources of Information:** contrary to the video-lead process described by literature [7, 61], participants expressed a preference to use their own recollection (P1, P3, P6), field notes (P2, P3, P6), voice recordings (P1, P4), conventional video (P3, P6), or photos (P1, P2, P3, P4, P6) as an initial resource in order to narrow down the *"...things we're interested in"* (P1) before engaging with 360° video, since there is *"... never enough time"* (P6) for analysis. This presented an additional friction for participants, since they had to find the relevant moments in 360° videos that were potentially stored on different devices (P1, P3) or properly documented (P2).

AG3 **Using 360° Video Beyond VDE:** the D360 tool is designed to support collaborative analysis between designers - however participants identified several ways they use 360° video outside of analysis. These included using D360 to familiarize new team members with a context during on-boarding (P1, P3, P5, P6) and using 360° imagery in reports generated for clients (P1, P4, P6). Additionally, participants discussed using the D360 system to aid in using evaluating design ideas or

(a) A mock-up of the "in-frame" annotation method described by P3.

(b) A mock-up of the map based overview for displaying annotations described by P2, P3, and P4.

Figure 7: Examples of two suggestions for improving D360 provided by participants, in this case used to represent an analysis of bicycle-traffic interactions.

prototypes. Examples given by participants include seeking to understand the impression of urban spaces that people had (P2, P3), asking others for feedback on sports performance (P4), or analyzing a recorded interaction with a prototype (P5).

6.3 RQ2: Utility of D360

Both when exploring the functionality in phases 3 and 4 as well as during the critique in phase 5, participants expressed that the D360 system provides additional utility for their use of 360° video. Participants were enthusiastic about the speed and simplicity of making annotations (P1, P2, P4, P6), especially compared with other 360° video software that *"...just doesn't let me note things down quickly"* (P1). The choice to integrate annotations with Miro was also seen as a way to streamline analysis (P1, P2, P3, P6) and aid with sharing insights from 360° video to other designers (P1, P2, P4, P6) and even as a way to present results to clients (P1, P6). Finally, participants pointed to the importance of quickly re-engaging with the 360° video (P1, P3, P4, P5, P6) in order to *"...really see the insight in the context"* (P5).

6.4 Additional Annotation Functionality

Throughout the sessions, participants provided a number of suggestions for improvements or additional functions (AF) for D360. Here we describe three such suggestions that focus on utility improvements[12]:

AF1 **Refined Annotation Input:** annotations created by D360 simply capture the viewpoint of the user in time, which provided enabled participants to quickly *"...add and annotations and just keep playing the video"* (P5). Participants did suggest more additional annotation functions such as drawing on the screenshots (P1, P2, P4, P6), using voice input (P3, P4), or providing quick numeric ratings (P2, P4). P3 expressed frustration that the button to create an annotation was outside the 360° video, and would prefer to directly add annotations to the 360° frame rather than the sidebar (Figure 7a).

AF2 **Meta-Frame Context:** while participants appreciated the addition of the mini-map in the annotation object (Figure 5b-B, P1, P2, P4, P6), there were several suggestions for context beyond single frame artifacts.
Some suggestions focused on improving the context around time. For example by providing an overview of the entire 360° video along with specific annotations (P1, P3), adding timestamps to annotations (P4), or providing a short animated sequence of frames before and after the annotation (P2, P6).
Another direction was to use a different dimension to provide context by overlaying annotations onto an interactive map of a city (P2, P3, P4 - Figure 7b) or a floor plan of the location the 360° video (P1, P5, P6). Another suggestion was to group annotations based on contextual elements (e.g., parks, supermarkets, factories) close to the camera in the moment of the annotation (P3, P4, P6).

AF3 **Greater Focus on Time:** Participants (P1, P3, P6) expressed frustration about not having enough control over the time of a 360° video video (e.g., frame by frame, playback speed, rewinding) in order to quickly navigate the whole video and then analyze specific moments in detail. When discussing how D360 places annotations in Miro on a timeline, P3 expressed how *"time is linear but not linear, you know?"* – some

[12]We thank the participants for their input on improving the UI of D360, however these are less transferable to other tools.

periods of 360° video have few annotations while other moments have many (P1, P3). Another element of non-linearity is the fact that some processes that are being redesigned are not recorded in the correct sequence (P6) or in multiple smaller 360° videos (P1, P3) – requiring the rearranging or merging of different 360° videos, potentially from multiple cameras (P1).

7 Discussion

In short, our results show that (1) the design goals used in the development of D360 reflect the experiences of designers (RQ1), and (2) designers are able to utilize D360 to conduct collaborative 360° video analysis (RQ2).

This means that D360 provides the utility necessary to help designers actually leverage the benefits of 360° video in VDE [32] opening the door to providing richer and more impactful insights [56] in the messy, iterative, and collaborative workflows of designers [7, 61].

Here we will discuss future directions for D360 or similar tools to further support the use of 360° video in design, expand on the concept of "linking" between different types of user research material, and how these kinds of tools could be useful beyond design. Then we discuss the limitations of this study and how future work can overcome them. Finally, we discuss how designers can actively engage with the important issue of privacy and 360° video.

7.1 Towards 360° Video as Designer Clay

Ylirisku and Buur [61] discuss the concept of video as "design clay", providing designers with material they can mold into another video by editing, composing, or adding new footage in order to demonstrate the importance of certain current interactions or envision future ones that are enabled by their design. While the design goals behind D360 focus on the analysis of 360° video for user research, participants highlighted the value of using D360 to share insights as an output. In this way D360 can provide a rudimentary ability for designers to engage with 360° video as clay - molding by creating a chain of annotations that together provide a new sequence. To enable designers to truly use 360° video as "clay", future tools should provide designers with a simple way to edit 360° videos - quickly creating a series of clips which can be assembled together to illustrate a specific concept and then "baked" into a 360° video provocation that can be shared and serve as a new object of analysis. Moving beyond editing existing videos, tools such as D360 could expand to help designers further leverage 360° video for envisioning, prototyping, and sharing future interactions. For example, by enabling them to quickly sketch 360° storyboards, such as those explored by Henrikson et al. [16], which could be overlaid on the real context using 360° videos. Another direction would be to dynamically link clips together in 360° viewing software, giving designers the ability to create "choose your own adventure" 360° experiences that can be used to highlight or explore possible futures. By providing flexible creative tools that integrate into the workflows used by designers, 360° video can truly become a clay that designers use throughout the design process, from analysis, to prototyping, to deliverables.

7.2 Linking More Than Video

One of the main mechanisms of the D360 system is the "backlinks" created between annotations and the 360° video itself. This helps reduce the complexity of returning to a moment of 360° video described by both our participants and literature [35], which in turn supports the iterative process of sense-making engaged in by designers [33]. Moreover, these links ensure that the rich visual context of 360° video is at hand even when the insights are represented in a simplified, static form – the annotation object – which also enables insights from 360° video to be integrated with user data from other sources. This blending of multiple sources of information in analysis is a core part of design ethnography [37] and is used by participants to reduce the amount of time required to analyze 360° video (6.2).

Here we propose two new kinds of linking to support designers when engaged with multi-modal user research (with or without 360° video) by improving switching between modalities (6.2) and discover similar moments [37, p.54] respectively.

(1) **Cross link:** connecting moments (actions, descriptions, results) from different sources of user research material using cross-modal information retrieval [19]. For example, linking a moment described in field notes with the 360° video at that moment as well as a reflection of the subject of the video in a post-hoc interview.
(2) **Automatic link:** connecting *similar* moments together to provide multi-modal recommendations [12] for user research material that helps expand the understanding of the analysis or suggesting connections between annotations that would lead to new dimensions of analysis.

To implement these two types of linking in an easy to use and scalable way requires tremendous technological effort. Fortunately, previous research has already explored how to search videos using text [26, 43] or even images of objects [50], which can form the technological basis of both cross and automatic linking.

Crucially, we apply the lens of "linking" to frame the technological development of these AI tools as ways to enable and enhance the messy, iterative, and constructivist process designers [33, 37] rather than automate the entire process. Future tool-kits can provide these links to designers, enabling them to engage with vast amounts of different types of user research material easily and switch between different media to discover insights at the level of detail that fits their analysis at the moment.

7.3 Uses of D360 Beyond Design

To reiterate, the design goals discussed in Section 3 are directly aimed to support the highly iterative and flexible nature of *Design* Ethnography [10, 33, 37]. However, these goals also overlap with needs in more conventional ethnographic uses of 360° video, such as the work of Vatanen et al. [56], as well as additional use cases beyond ethnography, based on input from participants as well as examples from literature, which we discuss here. One use case is education, where 360° video is leveraged to help students engage with environments they would otherwise not be able to access such as construction sites [24] or operating rooms [42]. One example discussed by [47] is how expert analysis of 360° video can support firefighter training. By having an expert annotate a 360° video

with "events of interest" and using the video with the annotations to explain the event and the reasoning behind the importance to students. Here, D360 would provide a number of advantages, by being able to create an annotation session, an expert instructor would be able to share all the events of interest in-situ with the 360° video, enabling students to review these sessions asynchronously. Students could also use D360 to self-annotate a 360° video, the results of which are collected in Miro, providing the teacher with an overview of the students work. The teacher could then analyze their students annotations and quickly bring up the 360° video and associated annotations when discussing with students. Another example of supporting education would be by providing authoring tools (Section 7.1), an easy to access viewer, and the ability to ask questions or make remarks via annotations, D360 could support the use of 360° video as a way to for students to engage with cultural heritage [48].

7.4 Limitations of the Study

As recommended by Ledo et al. [25], our study focused on understanding the *utility* of the D360 tool for the specific target group, rather than the *usability*. However, our study is limited by the number and diversity of designers we engaged with. The emergent nature of 360° video technology [32, 55] restricted the pool of experienced designers we could recruit. Engaging with a larger and more diverse group of designers in future studies could provide more nuanced insights into the utility of the various functions of D360, as well as uncover additional design considerations specific to individual contexts. Nonetheless, our participants' reflections on the utility of D360 align with the activities of VDE described by Ylirisku and Buur [61], suggesting that the utility of D360 is generalizable to the overall practice of 360° VDE.

Additionally, due to the novelty of D360, we evaluated it using a usage scenario and walk-through evaluation. While these evaluation methods do give insights into the utility of D360, to provide a deep understanding designers' use of D360 – what annotations they make, how they differ from those made with conventional video, and how D360 specifically supports their workflow – requires designers to be able to adopt D360 in the wild. By providing the tool to designers everywhere and seeing the behavior of users, future studies can truly understand the utility of D360 without participant bias [13] and in real world scenarios [25]. This would provide rich insights into the longitudinal evolution of the use of D360 (i.e., how its utility changes over the life of the design process, as hinted at in Figure 6) as well as how transferable the utility of D360 is to designers from other backgrounds, working in other domains, and with different experiences.

By providing D360 as an open and accessible tool, we aim to enable future longitudinal studies to understand and expand on how designers engage with 360° video and what utilities of D360 they do not use or are missing.

7.5 Privacy and 360° video

When it comes to user research, 360° video sits at a crossroads of flexibility and privacy. While the ability to capture the full visual context enables designers to reframe their analysis to fit "their interests as a designer" [61] even as those interests shift and evolve over course of a design project, it also strips the subjects (active subject *and* passive ones such as bystanders) of 360° video of agency in analysis as well as giving them no ability to hide from the camera [55] and forces them to share everything. This presents a major ethical concern for the use of 360° video by designers - how can they respect the privacy of users when the advantage of 360° video is contrary to principles like Data Minimization[13]. Based on this friction, we call on designers who use 360° video to consider five principles when working with 360° video from users:

(1) **Restrain:** only use 360° video when the additional context is beneficial.
(2) **Inform:** ensure subjects understand that 360° video truly captures both the full visual context and audio around the camera.
(3) **Empower:** give subjects the ability to censor specific moments or areas of the 360° video.
(4) **Restrict:** limit what is available outside the design team by only sharing relevant videos, limit access to relevant stakeholders, and anonymize as much as possible.
(5) **Forget:** limit the retention time for 360° videos of individuals to prevent it being used for purposes the subject did not consent to originally.

While these principles are pragmatic guidelines, the larger discussion of the ethics of 360° video remains an open question [55] that intersects with ethics in user research in general. As the ones engaged with understanding the user and their context, designers and researchers should actively engage with this question, both individually and collectively, both within and next to projects.

8 Conclusion

The D360 system was developed to support rapid, iterative, and collaborative analysis of 360° video so that designers can leverage the additional visual context it provides. The system consists of a 360° video viewing and annotation tool that exports the annotations to Miro - enabling designers to engage with 360° video and integrate it into their existing workflows.

Our evaluation of the D360 system through demonstration and walk-through with experienced designers demonstrated its utility. Designers are able to use the viewer to quickly explore and annotate 360° video on multiple devices. The Miro integration enables designers to use familiar tools for collaborative analysis and the annotations generated by D360 maintain the full 360° context. Finally, the meta-data collected by the D360 system helps keep track of attributing annotations to team members and provides "backlinks" that connect the annotations back to the complete 360° video. Our analysis also elicited addition functions important to designers engaged with 360° video: (1) lowering the threshold of interacting with 360° content, (2) mixing multiple sources of information into user context, and (3) using 360° video for on-boarding new team members or to evaluate design ideas.

Participants indicated limitations in relation to how D360 creates annotations as well as how to better represent the context and time of annotations. Furthermore, there are limits in the scale and scope of our study, in large part due to the novelty of 360° video in practice.

[13] Only collecting directly relevant personal information. https://www.edps.europa.eu/data-protection/data-protection/glossary/d_en#:~:text=Data%20minimization

In order to enable future work in addressing these limitations, we have open sourced D360 at https://github.com/WoMeijerPhD/d360-viewer. By providing D360 as an open and accessible tool, we enable designers to immediately start using 360° video – laying the groundwork for naturalistic case studies that can understand the utility of D360 across different design teams. This also opens up D360 for contexts beyond design such as education or crowd sourcing, and as a research tool for understanding the use of 360° video in those contexts.

In summary, the D360 system represents a significant advancement in the tools available for video design ethnography. It provides a simple, web-based interface for viewing and annotating 360° video that feeds into existing workflows. As open source tool, D360 enables designers to immediately adopt 360° video.

References

[1] Tanja Aitamurto, Andrea Stevenson Won, Sukolsak Sakshuwong, Byungdoo Kim, Yasamin Sadeghi, Krysten Stein, Peter G Royal, and Catherine Lynn Kircos. 2021. From FOMO to JOMO: Examining the Fear and Joy of Missing Out and Presence in a 360° Video Viewing Experience. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445183

[2] Dan Archer and Katharina Finger. 2018. Walking in another's virtual shoes: Do 360-degree video news stories generate empathy in viewers? (2018).

[3] EunSeo Bang and Caglar Yildirim. 2018. Virtually Empathetic?: Examining the Effects of Virtual Reality Storytelling on Empathy. In *Virtual, Augmented and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation*, Jessie Y.C. Chen and Gino Fragomeni (Eds.). Vol. 10909. Springer International Publishing, Cham, 290–298. https://doi.org/10.1007/978-3-319-91581-4_21 Series Title: Lecture Notes in Computer Science.

[4] Miguel Barreda-Ángeles, Sara Aleix-Guillaume, and Alexandre Pereda-Baños. 2020. An "Empathy Machine" or a "Just-for-the-Fun-of-It" Machine? Effects of Immersion in Nonfiction 360-Video Stories on Empathy and Enjoyment. *Cyberpsychology, Behavior, and Social Networking* 23, 10 (Oct. 2020), 683–688. https://doi.org/10.1089/cyber.2019.0665

[5] Hugh Beyer and Karen Holtzblatt. 1999. Contextual design. *Interactions* 6, 1 (Jan. 1999), 32–42. https://doi.org/10.1145/291224.291229

[6] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 18, 3 (July 2021), 328–352. https://doi.org/10.1080/14780887.2020.1769238

[7] Jacob Buur, Euan Fraser, Soila Oinonen, and Max Rolfstam. 2010. Ethnographic video as design specs. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction (OZCHI '10)*. Association for Computing Machinery, New York, NY, USA, 49–56. https://doi.org/10.1145/1952222.1952235

[8] Jacob Buur, Mads Vedel Jensen, and Tom Djajadiningrat. 2004. Hands-only scenarios and video action walls: novel methods for tangible user interaction design. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM, Cambridge MA USA, 185–192. https://doi.org/10.1145/1013115.1013141

[9] Jacob Buur and Astrid Soendergaard. 2000. Video card game: an augmented environment for user centred design discussions. In *Proceedings of DARE 2000 on Designing augmented reality environments (DARE '00)*. Association for Computing Machinery, New York, NY, USA, 63–69. https://doi.org/10.1145/354666.354673

[10] Andrew Crabtree, Mark Rouncefield, and Peter Tolmie. 2012. *Doing Design Ethnography*. Springer Science & Business Media. Google-Books-ID: Irm2KKegDjQC.

[11] Nigel Cross. 1982. Designerly ways of knowing. *Design Studies* 3, 4 (Oct. 1982), 221–227. https://doi.org/10.1016/0142-694X(82)90040-0

[12] Yashar Deldjoo, Markus Schedl, Paolo Cremonesi, and Gabriella Pasi. 2020. Recommender Systems Leveraging Multimedia Content. *ACM Comput. Surv.* 53, 5 (Sept. 2020), 106:1–106:38. https://doi.org/10.1145/3407190

[13] Nicola Dell, Vidya Vaidyanathan, Indrani Medhi, Edward Cutrell, and William Thies. 2012. "Yours is better!": participant response bias in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 1321–1330. https://doi.org/10.1145/2207676.2208589

[14] Claudia Eckert and Martin Stacey. 2000. Sources of inspiration: a language of design. *Design Studies* 21, 5 (Sept. 2000), 523–538. https://doi.org/10.1016/S0142-694X(00)00022-3

[15] Milene Guerreiro Gonçalves, Carlos Cardoso, and Petra Badke-Schaub. 2011. AROUND YOU: HOW DESIGNERS GET INSPIRED. *DS 68-7: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 7: Human Behaviour in Design, Lyngby/Copenhagen, Denmark, 15.-19.08.2011* (2011), 404–413. https://www.designsociety.org/publication/30695/AROUND+YOU%3A+HOW+DESIGNERS+GET+INSPIRED

[16] Rorik Henrikson, Bruno Araujo, Fanny Chevalier, Karan Singh, and Ravin Balakrishnan. 2016. Multi-Device Storyboards for Cinematic Narratives in VR. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, Tokyo Japan, 787–796. https://doi.org/10.1145/2984511.2984539

[17] Ann Heylighen and Andy Dong. 2019. To empathise or not to empathise? Empathy and its limits in design. *Design Studies* 65 (Nov. 2019), 107–124. https://doi.org/10.1016/j.destud.2019.10.007

[18] Tero Jokela, Jarno Ojala, and Kaisa Väänänen. 2019. How people use 360-degree cameras. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*. ACM, Pisa Italy, 1–10. https://doi.org/10.1145/3365610.3365645

[19] Parminder Kaur, Husanbir Singh Pannu, and Avleen Kaur Malhi. 2021. Comparative analysis on cross-modal information retrieval: A review. *Computer Science Review* 39 (Feb. 2021), 100336. https://doi.org/10.1016/j.cosrev.2020.100336

[20] Jan Koenderink and Andrea van Doorn. 2017. The Planispheric Optic Array. *i-Perception* 8, 3 (June 2017), 2041669517705388. https://doi.org/10.1177/2041669517705388 Publisher: SAGE Publications.

[21] Merlijn Kouprie and Froukje Sleeswijk Visser. 2009. A framework for empathy in design: stepping into and out of the user's life. *Journal of Engineering Design* 20, 5 (Oct. 2009), 437–448. https://doi.org/10.1080/09544820902875033

[22] Mark A. M. Kramer. 2022. Innovating Methods: Observing In Situ Technology Experiences Using Novel Visual Research Methods. *International Journal of Human–Computer Interaction* 38, 3 (Feb. 2022), 253–262. https://doi.org/10.1080/10447318.2021.1938388 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2021.1938388.

[23] Kartikaeya Kumar, Lev Poretski, Jiannan Li, and Anthony Tang. 2022. Tourgether360: Collaborative Exploration of 360° Videos using Pseudo-Spatial Navigation. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 546:1–546:27. https://doi.org/10.1145/3555604

[24] Noriyuki Kushiro, Koshiro Nishinaga, and Toshihiro Mega. 2021. Training Tool on Structured Knowledge for Risk Management with VR Technology. In *HCI International 2021 - Late Breaking Papers: Multimodality, eXtended Reality, and Artificial Intelligence*, Constantine Stephanidis, Masaaki Kurosu, Jessie Y. C. Chen, Gino Fragomeni, Norbert Streitz, Shin'ichi Konomi, Helmut Degen, and Stavroula Ntoa (Eds.). Springer International Publishing, Cham, 262–281. https://doi.org/10.1007/978-3-030-90963-5_21

[25] David Ledo, Steven Houben, Jo Vermeulen, Nicolai Marquardt, Lora Oehlberg, and Saul Greenberg. 2018. Evaluation Strategies for HCI Toolkit Research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–17. https://doi.org/10.1145/3173574.3173610

[26] Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. 2023. BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models. https://doi.org/10.48550/arXiv.2301.12597 arXiv:2301.12597 [cs].

[27] Qingchuan Li, Jiaxin Zhang, Xin Xie, and Yan Luximon. 2021. How Shared Online Whiteboard Supports Online Collaborative Design Activities: A Social Interaction Perspective. In *Advances in Creativity, Innovation, Entrepreneurship and Communication of Design*, Evangelos Markopoulos, Ravindra S. Goonetilleke, Amic G. Ho, and Yan Luximon (Eds.). Springer International Publishing, Cham, 285–293. https://doi.org/10.1007/978-3-030-80094-9_34

[28] Andrés Lucero, Peter Dalsgaard, Kim Halskov, and Jacob Buur. 2016. Designing with Cards. In *Collaboration in Creative Design*, Panos Markopoulos, Jean-Bernard Martens, Julian Malins, Karin Coninx, and Aggelos Liapis (Eds.). Springer International Publishing, Cham, 75–95. https://doi.org/10.1007/978-3-319-29155-0_5

[29] Tuuli Mattelmäki. 2008. *Design probes*. University of Art and Design, Helsinki. OCLC: 1086221371.

[30] Paul Bruce McIlvenny, Jacob Gorm Davidsen, Artúr Barnabás Kovács, Nicklas Haagh Christensen, and Stefan Hvarregaard Tanderup. 2021. AVA360VR: Annotate, Visualise, Analyse 360 Video in Virtual Reality. (2021). https://doi.org/10.5281/zenodo.5139349

[31] Wo Meijer, Jacky Bourgeois, Tilman Dingler, and Gerd Kortuem. 2025. Tangi: a Tool to Create Tangible Artifacts for Sharing Insights from 360° Video.. In *Proceedings of the Nineteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '25)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3689050.3704928

[32] Wo Meijer, Jacky Bourgeois, Wilhelm Frederik van der Vegte, and Gerd Kortuem. 2024. Sphere Window: Challenges and Opportunities of 360° Video in Collaborative Design Workshops.. In *Nordic Conference on Human-Computer Interaction (NordiCHI 2024)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3679318.3685407

[33] Francis Müller. 2021. *Design Ethnography: Epistemology and Methodology*. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-60396-0

[34] Daniel Neubauer, Verena Paepcke-Hjeltness, Pete Evans, Betsy Barnhart, and Tor Finseth. 2017. Experiencing Technology Enabled Empathy Mapping. *The Design Journal* 20, sup1 (July 2017), S4683–S4689. https://doi.org/10.1080/14606925.2017.1352966 Publisher: Routledge _eprint: https://doi.org/10.1080/14606925.2017.1352966.

[35] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. CollaVR: Collaborative In-Headset Review for VR Video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, Québec City QC Canada, 267–277. https://doi.org/10.1145/3126594.3126659

[36] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. Vremiere: In-Headset Virtual Reality Video Editing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 5428–5438. https://doi.org/10.1145/3025453.3025675

[37] Nicolas Nova. 2014. *Beyond design ethnography*. SHS Publishing.

[38] Dan R. Olsen. 2007. Evaluating user interface systems research. In *Proceedings of the 20th annual ACM symposium on User interface software and technology (UIST '07)*. Association for Computing Machinery, New York, NY, USA, 251–258. https://doi.org/10.1145/1294211.1294256

[39] Beth Pagett. 2023. Using a 360° Camera to Record Natural Dyeing Craft Practice. *FormAkademisk* 16, 4 (Sept. 2023). https://doi.org/10.7577/formakademisk.5412 Number: 4.

[40] Srishti Palani, David Ledo, George Fitzmaurice, and Fraser Anderson. 2022. "I don't want to feel like I'm working in a 1960s factory": The Practitioner Perspective on Creativity Support Tool Adoption. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–18. https://doi.org/10.1145/3491102.3501933

[41] Lucas Paulsen and Jacob Gorm Davidsen. 2025. CAVA360VR2: An Open-Source Tool for Collaborative Learning and Immersive Analysis in 360-Degree VR Environments. *ISLS 2025 proceedings* (June 2025). https://github.com/BigSoftVideo/CAVA360VR2.0/releases/tag/v2.0 Publisher: International Society of the Learning Sciences (ISLS).

[42] Johanna Pirker and Andreas Dengel. 2021. The Potential of 360° Virtual Reality Videos and Real VR for Education—A Literature Review. *IEEE Computer Graphics and Applications* 41, 4 (July 2021), 76–89. https://doi.org/10.1109/MCG.2021.3067999 Conference Name: IEEE Computer Graphics and Applications.

[43] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. https://doi.org/10.48550/arXiv.2103.00020 arXiv:2103.00020 [cs].

[44] Raf Ramakers, Fraser Anderson, Tovi Grossman, and George Fitzmaurice. 2016. RetroFab: A Design Tool for Retrofitting Physical Interfaces using Actuators, Sensors and 3D Printing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 409–419. https://doi.org/10.1145/2858036.2858485

[45] Emma Rose and Alison Cardinal. 2018. Participatory video methods in UX: sharing power with users to gain insights into everyday life. *Communication Design Quarterly* 6, 2 (Oct. 2018), 9–20. https://doi.org/10.1145/3282665.3282667

[46] Tony Salvador, Genevieve Bell, and Ken Anderson. 1999. Design Ethnography. *Design Management Journal (Former Series)* 10, 4 (1999), 35–41. https://doi.org/10.1111/j.1948-7169.1999.tb00274.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1948-7169.1999.tb00274.x.

[47] Ayush Sarkar, Anh Nguyen, Zhisheng Yan, and Klara Nahrstedt. 2022. A 360-Degree Video Analytics Service for In-Classroom Firefighter Training. In *2022 Workshop on Cyber Physical Systems for Emergency Response (CPS-ER)*. 13–18. https://doi.org/10.1109/CPS-ER56134.2022.00009

[48] Elmedin Selmanović, Selma Rizvic, Carlo Harvey, Dusanka Boskovic, Vedad Hulusic, Malek Chahin, and Sanda Sljivo. 2020. Improving Accessibility to Intangible Cultural Heritage Preservation Using Virtual Reality. *Journal on Computing and Cultural Heritage* 13, 2 (June 2020), 1–19. https://doi.org/10.1145/3377143

[49] Taufik Akbar Sitompul and Markus Wallmyr. 2019. Analyzing Online Videos: A Complement to Field Studies in Remote Locations. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernando Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis Zaphiris (Eds.). Vol. 11748. Springer International Publishing, Cham, 371–389. https://doi.org/10.1007/978-3-030-29387-1_21 Series Title: Lecture Notes in Computer Science.

[50] Josef Sivic and Andrew Zisserman. 2009. Efficient Visual Search of Videos Cast as Text Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 4 (April 2009), 591–606. https://doi.org/10.1109/TPAMI.2008.111 Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

[51] Susan Leigh Star and James R. Griesemer. 1989. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19, 3 (1989), 387–420. https://www.jstor.org/stable/285080 Publisher: Sage Publications, Ltd..

[52] Maik Stührenberg. 2014. Extending standoff annotation. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis (Eds.). European Language Resources Association (ELRA), Reykjavik, Iceland, 169–174. https://aclanthology.org/L14-1274/

[53] Antti Surma-aho and Katja Hölttä-Otto. 2022. Conceptualization and operationalization of empathy in design research. *Design Studies* 78 (Jan. 2022), 101075. https://doi.org/10.1016/j.destud.2021.101075

[54] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. 2020. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. ACM, Virtual Event USA, 182–195. https://doi.org/10.1145/3379337.3415827

[55] Naoya Tojo, Tomoko Oto, and Sumaru Niida. 2021. How Ethnographic Practices Are Reconfigured with 360-degree Cameras. 115–122. https://doi.org/10.5220/0010639000003060

[56] Anna Vatanen, Heidi Spets, Maarit Siromaa, Mirka Rauniomaa, and Tiina Keisanen. 2022. Experiences in Collecting 360° Video Data and Collaborating Remotely in Virtual Reality. *QuiViRR: Qualitative Video Research Reports* 3 (Sept. 2022), a0005–a0005. https://doi.org/10.54337/ojs.quivirr.v3.2022.a0005

[57] Mark Westmoreland, A. P. Pauwelussen, and Silke van Diemen. 2022. Kaleidoscopic Vision: Immersive Experiments in Maritime Worlds. *Entanglements* 5, 1/2 (2022), 50–70.

[58] Mark R Westmoreland. 2020. 360° Video. In *The Routledge international handbook of ethnographic film and video*. Routledge, 256–266.

[59] Jianxiong Xiao, Krista A. Ehinger, Aude Oliva, and Antonio Torralba. 2012. Recognizing scene viewpoint using panoramic place representation. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2695–2702. https://doi.org/10.1109/CVPR.2012.6247991 ISSN: 1063-6919.

[60] Shoma Yamaguchi, Nami Ogawa, and Takuji Narumi. 2021. Now I'm Not Afraid: Reducing Fear of Missing Out in 360° Videos on a Head-Mounted Display using a Panoramic Thumbnail. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 176–183. https://doi.org/10.1109/ISMAR52148.2021.00032 ISSN: 1554-7868.

[61] Salu Pekka Ylirisku and Jacob Buur. 2007. *Designing with Video: Focusing the user-centred design process*. Springer Science & Business Media. https://link.springer.com/book/10.1007/978-1-84628-961-3

A Usage Scenario

Based on our engagement with designers using 360° video before and after the expert evaluation (Section 5), we describe how an iterative Visual Design Exploration (VDE) process aimed at identifying automation opportunities in industrial kitchens could benefit from the capabilities of D360. This work is conducted by a team of three designers — Alex, Ben, and Chloe — with support from engineer David. This usage scenario demonstrates how D360 facilitates designers conducting VDE based on observations of designers using 360° video for a VDE process. The design task and clusters are illustrative and do not reflect the quality of insights generated in a real process.

Initial Annotation. Alex and Ben independently annotate 360° videos using the D360 viewer. Due to motion sickness, Alex prefers the desktop interface, while Ben opts for the immersive VR headset experience. To accelerate the process, they annotate different sets of videos from various kitchen types. Each creates initial action clusters for potential automation. Alex's clusters are based on perceived emotional responses to tasks, while Ben's are grounded in perceived automation difficulty.

Collaborative Sense-Making. Alex and Ben spend three hours in a meeting room aligning their annotations. Alex begins by sharing their overall video impressions, annotation method, and cluster formation. Ben follows suit. They then use Miro to merge overlapping clusters. However, Ben questions the relevance of the "creative complex cutting" cluster as their videos lacked such actions. Alex promptly demonstrates these moments using the D360 viewer. The

session concludes with Alex and Ben agreeing to refine their annotation process by mapping tasks based on the "mundane-novel" and "easy-complex" axes.

Onboarding Chloe. New designer Chloe joins the team and uses the Miro board to understand the VDE process thus far. The D360 viewer clarifies ambiguous annotations. Chloe observes that the "difficult to automate" cluster primarily contains early annotations by Ben. They suggest Alex and Ben revisit relevant videos to reassess automation difficulty.

Sharing with David. To incorporate expert input on automation feasibility, David joins the design team for a generative session to create three functional concepts. Alex, Ben, and Chloe share their evolving understanding of kitchen dynamics throughout the VDE process. David proposes creating a task timeline, which Chloe promptly implements using the D360 viewer and Miro. The session results in three automation concepts outlined on the Miro board. Post-session, David leverages the Miro board and D360 links to inform other automation engineers about the specific needs and goals of the three concepts. Two concepts are subsequently prototyped and evaluated in a test kitchen. For convenience, the evaluation is recorded using 360° cameras but cropped to a conventional field of view due to the designers' familiarity with the context.

Importance of D360 This example highlights D360's role in supporting the iterative nature of VDE, particularly its ability to facilitate rapid video revisiting and collaboration – functions that are not possible with current 360° video tools. Additionally, the example illustrates how, by integrating annotations into conventional tools (i.e., Miro), D360 enables design teams to use 360° video along with other sources of information in collaborative. Finally, the example underscores the value of screenshots and annotations for conceptualizing and sharing 360° video insights, while D360 enables seamless re-engagement with the full 360° view when needed.