American Psychological Association

Technology, Mind & Society
2021 Conference Proceedings

© 2021 The Author(s)
ISBN:

http://dx.doi.org/10.1037/xxxxxxx

An Ontology for Human-Centered Analysis and Design of Virtual Reality Conferencing

Maxi Heitmayer[1], Martha G. Russell[2], Saadi Lahlou[1], and Roy D. Pea[2]
[1]London School of Economics and Political Science
[2]Stanford University

Remote conferencing for collaboration is a fast-growing field. Virtual Reality applications are experiencing a surge in popularity as they afford a feeling of co-presence and 'togetherness' that video-conferencing cannot. But VR environments for conferencing are not fully mature yet in terms of space, and even less so in terms of the human activities they are supposed to facilitate. Moreover, for researchers, designers and practitioners from various backgrounds working on and in these spaces, a rigorous, shared language for conferencing activities in virtual environments to facilitate the exchange between different stakeholders does not exist as yet. In this paper, we therefore propose a simple, clear method for the development of a human-centered ontology for VR-conferencing, combining existing methodological protocols with Activity Theory and Installation Theory. We then present the first iteration of the ontology focusing specifically on human activity which we developed based on a recent 5-week conference with experts from academia and industry that took place in a 3D VR-environment.
Keywords: Virtual Reality, Ontology, Conferencing, IVC, Installation Theory, Activity Grid

Conferencing is essential to the life of professional and knowledge communities (Lave & Wenger, 1991; Wenger, 1998): their members gather, touch base, and often lay the foundation for future work. Disseminating new insights in the community, creating connections and trust between members, and building social capital are but a few activities that have gradually become supported by a set of institutions, rituals, formats, artefacts, and competences which we culturally now know as "professional conferences". Recently, remote conferencing via the Internet has developed because it suits the needs of increasingly globalized communities, accommodates for time pressure, budget cuts and environmental concerns, and because technological progress has improved the capacities and the quality of digital information communications. The recent COVID-19 pandemic has significantly accelerated this trend and has effectively forced communities to adopt a format of working and collaborating remotely. Beyond videoconferencing, which currently is the most popular format, we believe that Virtual Reality (VR) conferencing will see a massive surge in popularity because it can provide a feeling of co-presence and 'togetherness' that video-conferencing cannot.

The social practices that develop in early VR communities will become accepted social rules of interaction ("VRtiquette", compared to the current "netiquette" that rules online text conversations (Lahlou et al., 2021). VR-conferencing will therefore likely be a pioneering ground for the development of practices of sociability in virtual spaces. How can we describe what people do (and can do) as avatars while VR-conferencing, why they do it, and how they experience and value their actions and the following results? Such questions are not only informative for social science; they are also crucial for the stakeholders in charge of developing and running such events and spaces - designers, financers, and organizers.

VR-conferencing is in its infancy. While digital platforms are available, the suite of affordances and user fluency necessary for successful Installations for Virtual Conferencing (IVC; see below) is not fully mature yet, especially in terms of the social rules (Hollan & Stornetta, 1992; Lahlou et al., 2021). While researchers, designers, and developers work to improve these spaces, we observed in our first-hand experience with VR-conferencing that a common language for the behavioral science of VR-conferencing does not exist yet among the various stakeholders.

In this paper, we propose the development of an ontology to describe behaviors in virtual conferencing, to bridge the gap between design and psychological approaches; and we demonstrate a method for doing so. We develop the first iteration of this ontology based on our experiences from the Global Innovation Leadership Program (GILP) 2021 at Stanford University, a recent 5-week conference with academia and leading industry experts from four continents as participants. Given the constraints of this paper, we limit our description to one role (presenter), but that same approach can be extended to all roles in VR-conferencing. Overall, this provides researchers and practitioners with an actionable

The authors acknowledge the following people: We would like to thank Adelaide Dawes, Jason Wilmot, and all participants of the GILP2021.
Correspondence concerning this article should be addressed to Maxi Heitmayer. Email: m.a.heitmayer@lse.ac.uk

method to further develop ontologies for virtual environments with considerations of the human experience.

Rationale

An ontology creates a formal descriptive framework by establishing the classes, relationships and constraints that act on concepts and entities within a given system (Gruber, 1993; Musen, 1992). It defines what exists in a domain and describes the structure of that information; the rules that prescribe how a new category or entity is created, how attributes are defined, and how constraints are established. Ontologies, therefore, help researchers and practitioners to understand and analyze complex domains by providing an organizing structure for knowledge in the respective research field, also enabling knowledge sharing across fields (Gómez-Pérez, 1998; Gruninger & Fox, 1995; Noy & McGuinness, 2001; Uschold & Gruninger, 1996). They address the same scientific description problem as "ideal-types" (Schütz, 1996; Shils & Finch, 1949), but in a more formalized manner usable for design.

Four trends call for an ontology for virtual conferencing: workforce shifts accompanying the transition to a knowledge economy, requiring collaboration of team members who are not co-located; the integration of multiple computing desktops; the interoperability of mobile and immersive media (VR, AR, XR); and the near-term explosive growth of machine learning and big data applications for understanding the world. These trends are technology driven. However, their domain is human activity and therefore we need to go beyond technical details and describe the entities and behaviors in a way that makes sense *in terms of human activity*. By establishing an activity-based ontology for virtual conferencing that can be applied to both immersive virtual and hybrid 2D or 3D environments, scholars and practitioners may be able to accelerate innovations that cultivate engagement, motivation, understanding, performance for critical thinking, problem solving, and social emotional intelligence – key skills for the future and for collaboration. A clear, universally applicable, descriptive language can power a range of applications and enable technical and nontechnical designers of human-centered systems to define goals, model processes, and measure success.

Current vocabulary development for VR tends to focus on the space and the elements that populate it. An ontology for VR-conferencing, however, must not only encompass the technical elements of the virtual space, but also the real, subjective experiences users have. Unlike in natural settings, everything an actor can do has to be coded as a possibility in the computational system. Additionally, since virtual environments are relatively new spaces, the rules of behavior, both individual and social, must be clarified to users, and yet at times are still to be decided upon and codified by the community. Further, these rules must be reified in the system. Is it allowed to fly over or teleport into spaces? Can an actor hide her real identity? Who has the right to move someone's avatar into another space, or kick someone out of a session? Because the amount and nature of potential transactions that can be programmed is theoretically limitless, clear definitions of roles, statuses, and behaviors are required. Taking account of psychological and sociological insights into human behavior during ontology development will lead to a more effective categorization and organization of the human experience during VR-conferencing, benefitting researchers and practitioners alike, and users too.

Building the ontology of virtual conferencing will be a collective work in progress. Further work beyond the activity classifications of this paper will be needed to articulate motives, attitudes, and representations. This paper intends to start the process and to show a path. In the following, we discuss our approach combining standard methods for ontology design with Activity theory and Installation theory. We then present the first iteration of the ontology, discuss in more detail the actions of a conference presenter, and highlight areas for future work.

Related Work on Virtual Environments

Virtual reality, before its recent growth, has long been a topic in science-fiction novels, and early military and industry applications of virtual environments date back to the 1970s (Rheingold, 1992). A virtual environment is:

> *Any software-generated structure that is able to contain, or function as an environment for, software-generated objects and events, and human interactions with those objects and events.* (Brey, 2003)

Software-generated objects and entities in virtual environments are more than fictional objects; they are interactive and respond to manipulation by users, and they often feature various affordances for multi-sensory perception (Brey, 1999, 2003; Stanovsky, 2004). Moreover, while the space itself may be virtual, the institutions and rules in these spaces, as well as users' experiences, have real consequences (Fauville et al., 2020; Herrera et al., 2020). This is in part because objects, actions, and actors hold roles and properties which are socially reified beyond their physical make-up (a janitor, a five-pound-note, a wedding, or scoring a goal), and which can be ontologically reproduced in virtual environments through *functions* of the form "X counts as Y (in Context C)", and supplemented with *indicators* that make the functions explicit (a *Keep Out* sign, a staff uniform; Brey, 2003). Virtual environments, thus, are spaces *sui generis*, different from both fictitious and physical entities, and experiences in virtual environments have personal impact and are perceived as real (Bailenson & Beall, 2006; Fauville et al., 2020; van Loon et al., 2018).

Although conferences have been held in environments such as Second Life as early as 2009, empirical research on the use of VR spaces has so far mostly focused on designing VR training and learning spaces for specific applications such as engineering or construction (Edward et al., 2010; Pradhananga et al., 2020; Samarasinghe et al., 2019; Singh et al., 2021; Zhang & Chen, 2018), medical training (Cecil et al., 2017; Pappa & Papadopoulos, 2018; Parham et al., 2019), learning at school (Osti et al., 2020; Wang et al., 2009), and awareness of environmental or social issues (Markowitz et al., 2018). And while there are several ontologies for other aspects of virtual environments (e.g. Gutiérrez A. et al., 2007; Peña Pérez Negrón, Muñoz, & Lara López, 2020), there is to the best of our knowledge no formal ontology for the human experience in virtual conferencing.

Model for Analysis of Transactions in IVCs

As a minimum, an ontology of VR-conferencing must provide operational concepts to describe behaviors (what people do and can do), and values (why they act, how they experience their actions and their results) in the virtual space. Our ontology attempts to define these concepts in terms of human activities. To do so, we adopted an operational version of Activity theory (Nosulenko et al., 2005) and Installation theory (Lahlou, 2017) as a structure for understanding and categorizing the various actions, roles, rewards, competences, and affordances of the conferencing experience in what has been termed Installations for virtual conferencing (IVC; Lahlou et al., 2021).

Activity theory considers activity as an oriented trajectory from a given state to a consciously represented expected final state ('goal'), driven by an urge ('motive') to reach an internal state of satisfaction (Leontiev, 1978; Mironenko, 2013; Nosulenko & Samoylenko, 2009). We use Activity theory to break the trajectory of activity into a succession of elemental steps - small problems to be solved - each of which can be seen as reaching a local subgoal in the conditions given by the environment. At each step, several actors may be involved in different capacities with different motives (e.g., at the reception desk, two participants meet an organizer, one to get a badge and the other to be registered). Indeed, activity is subject-centric: performed from the perspective of the subject (Lomov, 1982; Nosulenko et al., 2005).

We use Installation theory to understand which other components are involved in each specific action. Installation theory states that at a given point of activity (e.g., attending a lecture) participant behavior is channeled by three layers of components: local affordances in the environment (e.g., seats, displays), embodied competences in the subjects (previous knowledge or skills for interpreting the situation), and social regulation (institutions and local rules). Note that embodied competences and rules are part and parcel of the Installation; its distributed architecture is incomplete without them. The combination of these three layers creates, by feed-forward and feed-back, a narrow tunnel of possible behaviors for the actor, which is why people behave as expected (Lahlou, 2017). Together, Activity theory and Installation theory enable us to describe what is done, by whom, why, and with what results.

An **action**, then, is a "consciously controlled move," in contrast to an operation, which is an automatic move occurring below the threshold of consciousness (Lahlou, 2017). Actions are the basic units of our ontology (e.g., enter room, respond to question).

A **transaction** is a set of interconnected, interdependent actions. Typically, a transaction will involve an action by an actor and actions by other actors in response to that action; take for example question and answer, greetings and salutations, and social exchange more generally (Blau, 1964). In a transaction, an actor acts in accordance with her role and is treated in accordance with her status. If the transaction produces value for other actors (e.g., by getting them closer to a goal or satisfying their motives) then it is considered satisficing (Simon, 1947). All actions in a virtual environment are behaviors that have been coded into the objects and avatars populating it; conversely, anything that should or may happen in a virtual environment must be coded into objects and avatars in order to be possible. Optimally, all transactions in IVCs should be satisficing.

An ontology that describes actions and transactions, as well as their experienced outcomes, will enable social scientists to describe and assess what happens in a VR-conference. For example, one can evaluate transaction satisfaction, compare with previous editions of the same VR-conference or with other conferences, identify possible improvements, and provide instructions for redesign in operational terms (actions that can be scripted and coded). Additionally, one could examine how many interactions have occurred between participants from different backgrounds. One could further determine the type of transactions, and which moments have sparked them. This could foster knowledge transmission, transdisciplinary inquiry, diversity, inclusion, and more.

Procedure for Ontology Development

Ontology design is a well-developed field in computer science, and several methods for the development of ontologies for virtual environments with the help of software packages have been described (De Moor, 2005; Dragoni et al., 2015; Pellens et al., 2005). Of particular note is a method that focuses on adding beliefs, desires, and intentions to artificial actors in virtual environments (Evertsz et al., 2009). We used these procedural recommendations to supplement our general approach for ontology development and applied our method to the context of the GILP2021.

We chose VR-conferencing as the domain of our ontology, and for this analysis limited the scope to the hybrid transactions of a) transmission of knowledge (content space) and b) the facilitation of social exchange (relational space) based on the theoretical work of Barron (2003). We began with a bottom-up deductive method, extracting from empirical to conceptual and used Activity theory and Installation theory to identify classes and elements for subsequent iterations of the ontology. We conducted three major steps of ontology development with multiple design iterations in each step. In the first step, a team of people involved in the 'provision' of the conferencing experience (organizers, facilitators, tech support) defined the various activities involved in VR-conferencing. As we encountered transactions, we recorded each actor involved and their roles and statuses as they emerged. We then followed each actor in their activity, breaking it down into successive steps with specific subgoals. In this step-by-step analysis, we filled in an activity grid, in which each line represented an action, and listed the conditions for performance of each action in the three layers provided by Installation theory (see Fig. 1; see also Appendix B). Conveniently, the components can be functionally defined through the behavior they enable at a finer level of detail. To vocalize a question in a plenary, for example, the actor must have a microphone for transmission (affordance), know how to initiate speech in the specific soft- and hardware setup of the IVC (embodied competence), and know the local rules of VRtiquette around politeness, timing, etc. (social regulation). This grid was used to feed the construction of the ontology, and we carried out three iterations of development in this step. In the second step, a group of participant observers joined the team, and

Actor/Role	Activity Sub-step	Activity ID	Action	Actor's motives	Installation components: **Affordances**	**Embodied Competences**	**Social Regulation**
Presenter	Preliminary to plenary	16	Navigate to podium	OP	MO, LO	WK, RN, FY, TI	PI, OR
Presenter	Preliminary to plenary	17	Talk to audience	SR, KE	AF, WC	SP	VQ, EC
Presenter	Plenary Presentation	18	Hear audience reactions	SR, KE	AF, WC	HE	VQ, EC
Presenter	Plenary Presentation	19	See audience reactions	SR, KE	VF, VD	SE	VQ, EC
Presenter	Plenary Presentation	20	Read questions from audience	SR, KE	TC	SE	VQ, EC
Presenter	Plenary Presentation	21	Hear comments from the audience	SR, KE	LC, WC	HE	VQ, EC
Presenter	Plenary Presentation	22	Invite members of audience to speak	SR, KE	LC, WC	SE, HE, SP, GE, DM	EC, OC, PI
Presenter	Plenary Presentation	27	Acknowledge audience reactions	SR	VF, AF, DC, VD, WC	SP, GE	PI, EC, OC, VQ, PR
Presenter	Plenary Presentation	23	Invite members of audience to change location	OP	VF, AF, DC, VD, WC	SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	24	Receive signal/communications from organizer	OP	VF, AF, DC, VD	SE, HE, SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	25	Give signal/communicate with organizer	OP	VF, AF, DC, VD, LC, TC	SE, HE, SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	26	Give signal/communicate with tech support	OP	VF, AF, DC, VD, TC, LC	SE, HE, SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	28	Leave Podium	OP	MO, LO	WK, RN, FY, TI	EC, OR, PI
Presenter	Post plenary	29	Invite/acknowledge/have conversation with participant	SR	VF, AF, VD	SE, HE, SP, DG, DP	EC, OC, PI, PR
Presenter	Post plenary	30	Invite others to join conversation	SR	VF, AF, VD	SE, HE, SP, DG, DP	EC, OC, PI, PR

Figure 1. ***An excerpt of a presenter's actions during a plenary session and the three layers of the IVC. The acronyms stand for specific components. e.g., Speak (SP): "Relating to the ability of IVC users to use their avatar to share their voice with other avatars" (see Appendix A for a Glossary and Appendix B for the full list).***

we revised the activity dimensions paying particular attention to the actors and their various roles and statuses, again developing the ontology over multiple rounds. In the third step, a smaller group of the initial team involved in organizing the conference worked on another iteration of the ontology focusing specifically on including insights from previous work on ontologies for virtual environments, as well as learning and collaborative work in virtual environments, and evaluated the final ontology against the goals set out at the beginning of the design process.

Proposed Ontology for VR-Conferencing

Four highest-order classes (nodes) build the spine of our ontology: 1) **actors**, in the form of roles and statuses; 2) **activity steps**, which are behaviors per se (e.g., raising hand, speaking); 3) the three **installation components** - layers of the IVC; 4) **values** (see Fig. 2). Classes 1, 2, and 3 can be defined by behaviors. While this may be obvious for class 2 (actions are goal-directed behaviors) and class 3 (components are conditions for behaviors), it is also valid for class 1 (roles and statuses). Indeed, a role is the set of behaviors that can be legitimately expected from an actor by others, while a status is the set of behaviors an actor can expect from others in a given situation (Lahlou, 2017; Stoetzel, 1963). Class 4, values, is not defined by observable actions and continues to be the object of much psychological research, under the constructs linked to motives, representations, attitudes, etc. Although most IVCs share some basic values, specific user communities will add or highlight certain elements (e.g., inclusion, economic value); the list of values in currency in each IVC is to be determined empirically (see below).

At the second level, we list five different roles and statuses that actors occupied during the conference in the first node (Audience member, Facilitator, Organizer, Presenter, Tech Support). For activities, we list here only three major zones of activity in IVCs (onboarding, presentation, networking) for the sake of brevity. The individual steps are comprised of many activity sub-steps and the respective behaviors distributed among them, which are not shown because of space constraints. Finally, we list the three layers of installation components, provided by Installation theory. We present only the third level for the third node of the ontology. For each of the three layers of the IVC (affordances, competences, rules), we list the various components that characterize the user experience during VR-conferencing based on the identified actors and activities. The length of this list is essential and reflects the multitude of elements that enable satisficing actions and transactions, as well as the requirements for coding the activity grid.

Due to space constraints, values are not developed beyond the first level in this ontology. For values to be added at a subsequent stage of ontology development, we suggest a conference-specific, grounded approach, using the motives, attitudes, and representations which organizers and participants empirically declare (e.g., building social capital; learning new methods; etc.). For example, in a brief teleconference feedback interview reflecting on the user experience in the IVC that was carried out with all participants after the conference (n=26), one participant noted that she runs her own company and tried to maximize the number of transferable skills she could take away to pass on to her employees to further her business (P12). On the contrary, another participant argued that the conference felt like "being back in school" and that he saw the conference as a welcome break from his daily job and an opportunity to develop his network (P14).

The ontology, in its current state, already provides an operational language for the human-centered analysis of VR-conferencing. Figure 1 shows a brief excerpt from an activity grid detailing the actions a presenter had to go through to be able to deliver her presentation during the GILP2021, from navigating the virtual conferencing hall, over monitoring audience reactions, to discussing her work with individual attendees after the presentation, and how the various elements identified in the three layers of the IVC influenced how these actions played out (see Glossary in Appendix A). For example, when inviting an audience member to speak (Fig. 1, step 23) the local and the wide audio channel (affordances), her ability to speak, hear, see, make gestures and manipulate objects with her avatar, and her understanding of the agenda and time-keeping skills (embodied competences), as well as her understanding of the organizational and the local event culture and her personal interest (social regulation), influence the success of the activity and the nature of the outcome. Together with the complementary action of responding to the presenter for audience members, these form the VR-conferencing transaction "asking a question during the plenary". The proposed ontology of virtual conferencing provides a set of shared concepts and terms

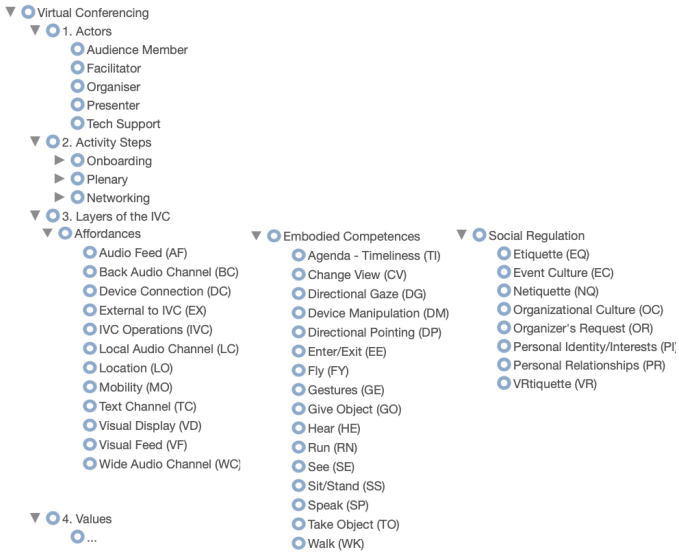

Figure 2. ***Spine of the ontology for VR-conferencing with a detailed view of the elements of three layers of the IVC.***

that will aid in understanding the unique human experiences of each virtual conference - its context, objectives, and participants. Due to spatial constraints of this paper, we have limited our description to details of the presenter role. Even so, this process can describe all relevant actors, objects, and transactions for VR-conferencing with a human-centered approach. This system is simple, reliable, and obviates a need for large amounts of expertise since it is based on filling in an activity grid with empirical observations of conferencing actions: in practice, doing this work has been experienced without difficulty by participants of diverse backgrounds (psychologists, designers, engineers, social scientists, and admin support). We anticipate that this first step towards a comprehensive language for virtual conferencing will facilitate meaningful dialogue in teams of designers, developers, financers, and organizers of VR-conferences, enabling engagement in a wide array of virtual interactions to improve user experience and productivity. This ontology contributes to a foundational language for sharing across scholarly disciplines engaged in studying human communicative behavior in virtual environments, and it creates a linguistic bridge between technical, operational, and scholarly communities developing, promoting, and studying the domain of virtual conferencing. Of course, it is work in progress.

Conclusion and Outlook

To allow conferences in IVCs to realize their full potential, a synergy of affordances, embodied competences, and social regulation is required. This paper describes a simple, reliable, and practical method for developing an ontology for VR-conferencing, focusing on users' activities. It also provides the first iteration of the ontology and, thus, a first step towards a shared language of collaborative activity in virtual environments that can streamline and improve the objects and transactions that create and govern the IVC. It will further enable new understandings of collaboration and communication in human experiences, and new modes of inquiry, including, but not limited to:

- Designing better and more user-friendly IVCs
- Assessing, measuring, predicting user experience
- Methods and metrics to understand the interactivity between cognitive, emotional, and physical states
- Conducting and sharing research on VR-conferencing, especially psychological and behavioral; and transferring findings to designers with minimal loss in translation.

References

Bailenson, J. N., & Beall, A. C. (2006). Transformed social interaction: Exploring the digital plasticity of avatars. In R. Schroeder & A. Axelsson (Eds.), *Avatars at work and play: Collaboration and interaction in shared virtual environments* (pp. 1–16). Springer Verlag. https://doi.org/10.1007/1-4020-3898-4_1

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*. https://doi.org/10.1207/S15327809JLS1203_1

Blau, P. M. (1964). *Exchange and power in social life*. New York:Wiley.

Brey, P. (1999). The ethics of representation and action in virtual reality. *Ethics and information technology*. https://doi.org/10.1023/A:1010069907461

Brey, P. (2003). The social ontology of virtual environments. *American Journal of Economics and Sociology*, *62*(1), 269–282. https://doi.org/10.1111/1536-7150.t01-1-00011

Cecil, J., Kumar, M. B. B. R., Gupta, A., Pirela-Cruz, M., Chan-Tin, E., & Yu, J. (2017). Development of a virtual reality based simulation environment for orthopedic surgical training. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-55961-2_21

De Moor, A. (2005). Towards ontology-guided design of learning information systems. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *3762 LNCS*, 1190–1199. https://doi.org/10.1007/11575863_142

Dragoni, M., Ghidini, C., Busetta, P., Fruet, M., & Pedrotti, M. (2015). Using ontologies for modeling virtual reality scenarios. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-18818-8_35

Edward, L., Amokrane, K., Lourdeaux, D., & Barthès, J. P. (2010). An ontology for managing a virtual environment for risk prevention. *Proceedings - 1st International Conference on Integrated Intelligent Computing, ICIIC 2010*, 62–67. https://doi.org/10.1109/ICIIC.2010.17

Evertsz, R., Pedrotti, M., Busetta, P., Acar, H., & Riller, F. E. (2009). Populating VBS2 with realistic virtual actors. *18th Annual Conference on Behavior Representation in Modeling and Simulation 2009, BRiMS 2009*, *April*, 1–8.

Fauville, G., Queiroz, A. C. M., & Bailenson, J. N. (2020). Virtual reality as a promising tool to promote climate change awareness. In *Technology and Health*. https://doi.org/10.1016/b978-0-12-816958-2.00005-8

Gómez-Pérez, A. (1998). Knowledge sharing and reuse. In J. Liebowitz (Ed.), *The Handbook of Applied Expert Systems*. CRC Press. https://doi.org/10.1201/9780138736654-10

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199-220.

Gruninger, M., & Fox, M. S. (1995). Methodology for the design and evaluation of ontologies. *International Joint Conference on Artificial Intelligence (IJCAI95), Workshop on Basic Ontological Issues in Knowledge Sharing*.

Gutiérrez A., M., García-Rojas, A., Thalmann, D., Vexo, F., Moccozet, L.,

Magnenat-Thalmann, N., Mortara, M., & Spagnuolo, M. (2007). An ontology of virtual humans : Incorporating semantics into human shapes. *Visual Computer*. https://doi.org/10.1007/s00371-006-0093-4

Herrera, F., Oh, S. Y., & Bailenson, J. N. (2020). Effect of behavioral realism on social interactions inside collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *27*(2), 163–182. https://doi.org/10.1162/PRES_a_00324

Hollan, J., & Stornetta, S. (1992). Beyond being there. In P. Bauersfeld, J. Bennett, & G. Lynch (Eds.), *Conference on Human Factors in Computing Systems - Proceedings* (pp. 119–125). Association for Computing Machinery, NY, USA. https://doi.org/10.1145/142750.142769

Lahlou, S. (2017). *Installation theory: the societal construction and regulation of behaviour*. Cambridge University Press. https://doi.org/10.1017/9781316480922

Lahlou, S., Pea, R., Heitmayer, M., Russell, M. G., Schimmelpfennig, R., Yamin, P., Everri, M., Cordelois, A., & Dawes, A. P. (2021). Are we 'Beyond being there' yet?: Towards better interweaving epistemic and social aspects of virtual reality conferencing. *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '21 Extended Abstracts)*. https://doi.org/10.1145/3411763.3451579

Lave, J., & Wenger, E. C. (1991). *Situated learning : legitimate peripheral participation*. Cambridge University Press.

Leontiev, A. N. (1978). *Activity, consciousness, and personality*. Prentice-Hall.

Lomov, B. F. (1982). The problem of activity in psychology. *Soviet Psychology*, *21*, 55–91.

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive virtual reality field trips facilitate learning about climate change. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.02364

Mironenko, I. A. (2013). Concerning interpretations of activity theory. *Integrative Psychological and Behavioral Science*, *47*(3), 376–393. https://doi.org/10.1007/s12124-013-9231-5

Musen, M. A. (1992). Dimensions of knowledge sharing and reuse. *Computers and Biomedical Research*, *25*(5), 435–467. https://doi.org/10.1016/0010-4809(92)90003-S

Nosulenko, V. N., Barabanshikov, V. A., Brushlinsky, A. V., & Rabardel, P. (2005). Man–technology interaction: Some of the Russian approaches. *Theoretical Issues in Ergonomics Science*, *6*(5), 359–383. https://doi.org/10.1080/14639220500070051

Nosulenko, V. N., & Samoylenko, E. S. (2009). Psychological methods for the study of augmented environments. In S. Lahlou (Ed.), *Designing user friendly augmented work environments. From meeting rooms to digital collaborative spaces* (pp. 213–236). Springer. https://doi.org/10.1007/978-1-8400-098-8

Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. *Stanford Knowledge Systems Laboratory*, 25. https://doi.org/10.1016/j.artmed.2004.01.014

Osti, F., de Amicis, R., Sanchez, C. A., Tilt, A. B., Prather, E., & Liverani, A. (2020). A VR training system for learning and skills development for construction workers. *Virtual Reality*. https://doi.org/10.1007/s10055-020-00470-6

Pappa, D., & Papadopoulos, H. (2018). Designing a prototype training environment for physiotherapists building on advanced gaming technologies. *Proceedings of the European Conference on E-Learning, ECEL*.

Parham, G., Bing, E. G., Cuevas, A., Fisher, B., Skinner, J., Mwanahamuntu, M., & Sullivan, R. (2019). Creating a low-cost virtual reality surgical simulation to increase surgical oncology capacity and capability. *Ecancermedicalscience*. https://doi.org/10.3332/ecancer.2019.910

Pellens, B., De Troyer, O., Bille, W., Kleinermann, F., & Romero, R. (2005). An ontology-driven approach for modeling behavior in virtual environments. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/11575863_145

Peña Pérez Negrón, A., Muñoz, E., & Lara López, G. (2020). A model for nonverbal interaction cues in collaborative virtual environments. *Virtual Reality*, *24*(4), 605–618. https://doi.org/10.1007/s10055-019-00421-w

Pradhananga, P., Elzomor, M., Santi, G., & Zhang, L. (2020). Fostering virtual reality environments to advance construction and engineering students' interpersonal skills. *ASEE Annual Conference and Exposition, Conference Proceedings*. https://doi.org/10.18260/1-2--34686

Samarasinghe, D. A. S., Latif, S. A., & Baghaei, N. (2019). Virtual reality models for promoting learners engagement in construction studies. *IEEE Global Engineering Education Conference, EDUCON*. https://doi.org/10.1109/EDUCON.2019.8725120

Rheingold, H. (1992). *Virtual reality:The revolutionary technology of computer-generated artificial worlds and how it promises to transform society]*. New York: Simon & Schuster.

Schütz, A. (1996). *Collected Papers, Volume IV* (H. Wagner, G. Psathas, & F. Kersten (eds.)). Springer-Science+Business Media, BV. https://doi.org/10.1007/978-94-017-1077-0

Shils, E. A., & Finch, H. A. (Eds.). (1949). *Max Weber on the methodology of the social sciences*. New York: Free Press.

Simon, H. A. (1947). *Administrative Behavior*. Macmillan.

Singh, G., Mantri, A., Sharma, O., & Kaur, R. (2021). Virtual reality learning environment for enhancing electronics engineering laboratory experience. *Computer Applications in Engineering Education*. https://doi.org/10.1002/cae.22333

Stanovsky, D. (2004). Virtual Reality. In L. Floridi (Ed.), *The Blackwell Guide to the Philosophy of Computing and Information* (pp. 167–177). Oxford: Blackwell Publishing.

Stoetzel, J. (1963). *La psychologie sociale*. Flammarion.

Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *Knowledge Engineering Review*. https://doi.org/10.1017/s0269888900007797

van Loon, A., Bailenson, J., Zaki, J., Bostick, J., & Willer, R. (2018). Virtual reality perspective-taking increases cognitive empathy for specific others. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0202442

Wang, X., Zhou, M., & Fan, Y. (2009). Building VR learning environment: An ontology based approach. *Proceedings of the 1st International Workshop on Education Technology and Computer Science, ETCS 2009*, *3*, 160–165. https://doi.org/10.1109/ETCS.2009.561

Wenger, E. C. (1998). Communities of practice: Learning as a social system. *Systems Thinker*, *9*(5 (June/July)), 1–5.

Zhang, C., & Chen, A. B. (2018). VR-based teaching and learning tool for building design and construction. *EUCEET 2018 - 4th International Conference on Civil Engineering Education: Challenges for the Third Millennium*.

Appendix

A) Glossary

This Glossary provides descriptions of the Installation components in the three layers presented in Figure 1, 2, and B2.

Affordances

Audio feed (AF): Relating to the audio feed of the IVC, e.g., hardware setup allows user to access audio information provided by the IVC.

Back audio channel (BC): Relating to the back audio channel either internal or external to the IVC, e.g., WhatsApp group chat for organizers or working groups.

Device Connection (DC): Relating to the connective capacities of the machine used to access the IVC, e.g., laptop, tablet, desktop computer.

External to IVC (EX): Relating to external affordances used while conferencing in the IVC, e.g., Email, telephone.

IVC Operations (IVC): Relating to the demonstration and administrative facilities of the IVC, e.g., landing spot, onboarding tutorial.

Local audio channel (LC): Relating to the local audio channel, e.g., in the specific room or area of the IVC.

Location (LO): Relating to a specific location in the IVC, e.g., corridor, auditorium, breakout room.

Mobility (MO): Relating to the capacities of movement of the avatar in an IVC, e.g., traversing space, walking or running speed.

Text channel (TC): Relating to the text channel of the IVC, e.g., local or waiting room chat.

Visual display (VD): Relating to Visual Displays in the IVC, e.g., presentation theater screens, help desk display, virtual machines.

Visual feed (VF): Relating to the audio feed of the IVC, e.g., hardware setup allows user to access visual information provided by the IVC.

Wide audio channel (WC): Relating to the wide audio channel, e.g., for general announcements in the IVC or in areas without local audio zones.

Embodied Competences

Agenda - Timeliness (TI): Relating to the ability of IVC users to adhere to timings and follow guidelines for specific sessions and agenda items.

Change View (CV): Relating to the ability of IVC users to change view.

Directional Gaze (DG): Relating to the ability of IVC users to do directional gaze.

Device Manipulation (DM): Relating to the ability of IVC users to operate the computer and A/V hardware used to access the IVC.

Directional Pointing (DP): Relating to the ability of IVC users to point with their avatar.

Enter/Exit (EE): Relating to the ability of IVC users to enter and exit spaces with their avatar.

Fly (FY): Relating to the ability of IVC users to fly with their avatar.

Gestures (GE): Relating to the ability of IVC users to display gestures and emotes with their avatar.

Give object (GO): Relating to the ability of IVC users to give objects to other avatars.

Hear (HE): Relating to the ability of IVC users to hear and locate sound their avatar encounters.

Run (RN): Relating to the ability of IVC users to run with their avatar.

See (SE): Relating to the ability of IVC users to see and experience visual information their avatar encounters.

Sit/Stand (SS): Relating to the ability of IVC users to sit down and stand up with their avatar.

Speak (SP): Relating to the ability of IVC users to use their avatar to share their voice with other avatars.

Take object (TO): Relating to the ability of IVC users to pick up objects from the environments or take objects from other avatars.

Walk (WK): Relating to the ability of IVC users to walk with their avatar.

Social Regulation

Etiquette (EQ): The general rules of conduct that govern social interactions, e.g., raise your hand before speaking up.

Event culture (EC): The specific spirit and culture of a work or social event, e.g., at GILP2021, we use first names and directly engage people even if we have not had contact with them before.

Netiquette (NQ): The general rules of conduct that govern social interactions on the internet, e.g., mute yourself after speaking.

Organizational culture (OC): The specific spirit and culture of an organization or institution.

Organizer's request (OR): Relating to a specific request made by the organizer, e.g., provide a 300-word bio to be included in the conference program.

Personal Identity or Interests (PI): Relating to the personal interests and identities of IVC users, and how these govern their actions and decisions.

Personal relationships (PR): Relating to the personal relationships IVC users have with other participants, and how these influence their actions and decisions.

VRtiquette (VQ): The general rules of conduct that govern social interactions in VR environments, e.g., use clapping and cheering gestures to provide audience feedback, park avatar or mark as away from keyboard (afk).

B) Additional Materials from the Activity analysis

Figure B1 gives a broad overview of the activities in which a presenter engages while conferencing in IVCs. Figure B2 provides the full list of actions associated with giving a plenary presentation in an IVC.

Figure B1. ***A list of activities a presenter engages in while conferencing (adapted from Lahlou et al., 2021)***

	Task	Activities
1.	Awareness	Getting info/invitation: there is a conference at this time and place, about these topics.
2.	Timetabling	Planning, booking, solving authorization issues (clearance from organization, funding).
3.	Preparation	Writing, reviewing, and editing paper, coordination with organizers and tech support.
4.	Onboarding	Travelling or exploration of the digital platform; may include getting and testing the display installation (this phase is a bit different from in-person for IVCs because tests can start early).
5.	Orientation	Creating more detailed activity plans once more aware of resources on site.
6.	Presentation	Speaker to audience, data display, moderation, speaker interaction in panels.
7.	Audience Interaction	A good presentation usually includes interaction with the audience, e.g., Q&A, laughter, and applause.
8.	Breaks and transitions	Social interaction, transfers between sessions, networking, physiological pause, keeping in touch with "normal work".
9.	Workshops	N to N participant interaction, producing collective outputs for proceedings.
10.	Visits and Socializing	Visits, tours, meals, and other activities and opportunities to meet like-minded people and to network.
11.	Disembarking	Changing settings, uninstalling software, and rearranging workstations.
12.	Follow-up	Storage/retrieval of material and contacts from the conference, sharing of material produced based on the conference for later publication.

Figure B2. ***A presenter's actions during a plenary session and the three layers of the IVC***.

Actor/Role	Activity Sub-step	Activity ID	Action	Actor's motives	Installation components: **Affordances**	**Embodied Competences**	**Social Regulation**
Presenter	Preliminary to plenary	01	Provide title to talk for organizers	KE	EX	.	OC, OR
Presenter	Preliminary to plenary	02	Provide bio and image to organizers	KE	EX	.	OC, OR
Presenter	Preliminary to plenary	03	Provide brief description or video snip to organizers	KE	EX	.	OC, OR
Presenter	Preliminary to plenary	04	Prepare/modify content	KE	EX	.	OR, PI
Presenter	Preliminary to plenary	05	Review session logistics	KE, OP	EX	.	PI
Presenter	Preliminary to plenary	06	Enter IVC	OP	IVC	EE	OR
Presenter	Preliminary to plenary	07	Complete tutorial	OP	IVC	EE	OR, PI
Presenter	Preliminary to plenary	08	Test AV	OP	VF, AF, DC, VD, LC	DM	OR
Presenter	Preliminary to plenary	09	Enter IVC in avatar	OP	IVC	EE	VR, EC
Presenter	Preliminary to plenary	10	Navigate to auditorium	OP	MO, LO	WK, RN, FY, TI	VR, EC, PI
Presenter	Preliminary to plenary	11	Connect tested AV	OP	VF	DM	OR
Presenter	Preliminary to plenary	12	Communicate with tech support	OP	LC	SE, HE, SP	VR
Presenter	Preliminary to plenary	13	Adjust AV	OP	VF, AF, DC, VD	DM	OR, PI
Presenter	Preliminary to plenary	14	Communicate with organizers	SR, OP	LC	SE, HE, SP	VR, OC, PI
Presenter	Preliminary to plenary	14	Learn number/identity/location of attendees	SR, OP	VF, AF, DC, VD	SE, HE, SP	OR, PI
Presenter	Preliminary to plenary	15	Control AV used in presentation	OP	DC	DM	OR, EC, OC, PI
Presenter	Preliminary to plenary	16	Navigate to podium	OP	MO, LO	WK, RN, FY, TI	PI, OR
Presenter	Preliminary to plenary	17	Talk to audience	SR, KE	AF, WC	SP	VQ, EC
Presenter	Plenary Presentation	18	Hear audience reactions	SR, KE	AF, WC	HE	VQ, EC
Presenter	Plenary Presentation	19	See audience reactions	SR, KE	VF, VD	SE	VQ, EC
Presenter	Plenary Presentation	20	Read questions from audience	SR, KE	TC	SE	VQ, EC
Presenter	Plenary Presentation	21	Hear comments from the audience	SR, KE	LC, WC	HE	VQ, EC
Presenter	Plenary Presentation	22	Invite members of audience to speak	SR, KE	LC, WC	SE, HE, SP, GE, DM	EC, OC, PI
Presenter	Plenary Presentation	27	Acknowledge audience reactions	SR	VF, AF, DC, VD, WC	SP, GE	PI, EC, OC, VQ, PR
Presenter	Plenary Presentation	23	Invite members of audience to change location	OP	VF, AF, DC, VD, WC	SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	24	Receive signal/communications from organizer	OP	VF, AF, DC, VD	SE, HE, SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	25	Give signal/communicate with organizer	OP	VF, AF, DC, VD, LC, TC	SE, HE, SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	26	Give signal/communicate with tech support	OP	VF, AF, DC, VD, TC, LC	SE, HE, SP	PI, EC, OC, VQ
Presenter	Plenary Presentation	28	Leave Podium	OP	MO, LO	WK, RN, FY, TI	EC, OR, PI
Presenter	Post plenary	29	Invite/acknowledge/have conversation with participant	SR	VF, AF, VD	SE, HE, SP, DG, DP	EC, OC, PI, PR
Presenter	Post plenary	30	Invite others to join conversation	SR	VF, AF, VD	SE, HE, SP, DG, DP	EC, OC, PI, PR
Presenter	Post plenary	31	End conversation with participant	SR	VF, AF, VD	SE, HE, SP, DG, DP	EC, OC, PI, PR
Presenter	Post plenary	32	Exchange digital information with audience member	SR, KE, PR	DC	DM	EC, OC, PR
Presenter	Personal needs & comfort	33	Communicate with back office (not VR)	PY, RE	.	.	OC, EC, PI, EQ, NQ
Presenter	Personal needs & comfort	34	Receive urgent text (not in VR)	PY, RE	.	.	OC, EC, PI, EQ, NQ
Presenter	Personal needs & comfort	36	Put ringing device on mute - not in VR	PY, RE	.	.	OC, EC, PI, EQ, NQ
Presenter	Personal needs & comfort	37	Answer urgent non-VR call	PY, RE	.	.	OC, EC, PI, EQ, NQ
Presenter	Personal needs & comfort	35	Attend to personal comfort - bathroom, drink	CO	.	.	EC, PI