Stefan Weinzierl and Steffen Lepa

On the Epistemic Potential of Virtual Realities for the Historical Sciences. A Methodological Framework

Abstract: Virtual or augmented audio-visual environments can be employed not only for the impartment of knowledge to a wider audience, but also for the generation of knowledge within the historical sciences. In this context, the transformation of numerical models of historical circumstances into an immediate sensual experience may be used both in an exploratory manner as well as for testing specific hypotheses through subjective perceptual analysis. As with any other empirical approach, the new insights provided can be biased on different levels. In order to make virtual realities (VRs) a valuable tool for research, traditional quality criteria for empirical research need to be adapted to the specific setting created by observational fieldwork 'within' 3D audio-visual computer simulations. Two major causes for degradation in the credibility of VR-based historical research are related to the relationship between simulation and historical 'reality' as well as to the human agents experiencing the simulated environments and the conclusions drawn from their subjective impressions. Hence, our contribution attempts to outline procedures and methods for estimating and comparing the ecological validity of virtual environments as well as the level of intersubjectivity regarding the inferences drawn during and after experiencing them. For this purpose, we suggest to synthesize existing ideas and procedures originating from virtual reality research, media psychology, communication science and ethnology.

Keywords: Virtual Reality, Augmented Reality, Virtual Acoustic Reality, Virtual Historic Environment, Virtual Archaeology, Cyber Ethnography.

Immersive media environments have been used for quite some time for the *impartment* of historical knowledge to a wider audience (Mikropoulos/Natsis 2011). They have a long tradition reaching from panoramic paintings (Oettermann 1980) and diorama reconstructions of historical events (Gernsheim/Gernsheim 1969) to virtual environments using recent 3D technologies in museums (Stogner 2009, Bearman 2011), or cultural heritage projects (Bogdanovych *et al.* 2010, Pujol/Champion 2012). Examples include the virtual reconstruction of urban space and daily life in 13th-century Bologna, key historical theaters of

DOI 10.1515/9783110497656-004

18th-century Venice (Lercari 2016), or of audio-visual media installations such as the *Poème électronique* and the Philips pavilion by Le Corbusier and Edgard Varèse at the Brussels world fair in 1958 (Lombardo *et al*. 2009).

Less frequent are attempts to *generate new knowledge* from the virtual reconstruction of historical scenarios, and much greater are the reservations within the scientific community against this kind of scientific practice. These reservations may be differently motivated in each individual case. In their core, however, they are related a) to doubts regarding the credibility of virtual environments as a form of scientific evidence; b) to the question of how new knowledge may emerge from virtual environments, which are always based on already existing historical evidence; and c) what kind of role virtual or augmented realities could play within an epistemological concept of historical research. With the current text, we would like to sketch out a methodological framework that will attempt to provide answers to these questions.

This framework will be exemplified by a concrete problem drawn from the authors' own research. The open question concerned a) the maximum audience size reached by a speaker on the ancient Forum Romanum between ca. 500 BC and ca. 500 AD; b) how the size of this audience changed with the several modifications and relocations of the Rostra; and c) whether these modifications were carried out for reasons of acoustics, or rather for the political and symbolic motives usually invoked in the historiography of the Roman empire. This concrete problem was part of a larger project dedicated to the digital reconstruction of the ancient Forum Romanum[1] and worked on by a team of archeologists and acousticians.

In the following outline for a methodological approach towards the use of virtual realities for historical research, we will restrict ourselves to the acoustic domain, which was particularly important for the given problem, although all technical and methodological stages have an almost one-to-one equivalent in the visual domain. Since many of the methodological problems are closely related to the technical process of generating a virtual environment, we will first give a brief outline of binaural synthesis as the chosen approach for virtual acoustic reality (Fig. 2).

1 See http://www.digitales-forum-romanum.de/, accessed on 20 September 2016.

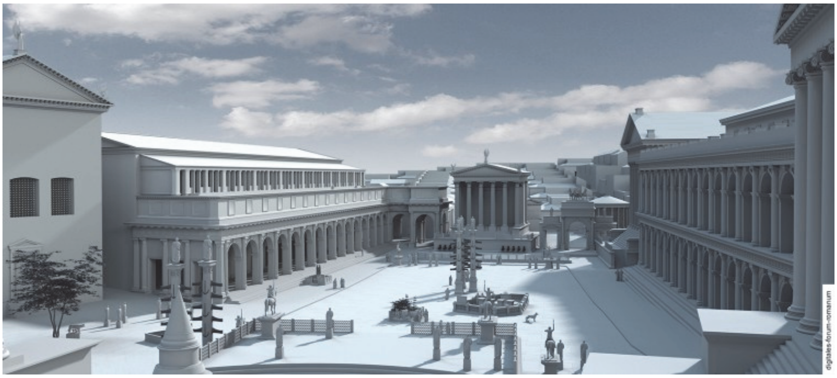

Fig. 1: Digital reconstruction of the ancient Forum Romanum. The model represents the conditions in 14 AD, the year of Augustus's death. View from southeast.

1. Virtual acoustic environments

Binaural synthesis is one approach for the generation of virtual acoustic environments. It is the acoustic equivalent of stereoscopic displays, in which one image is generated for each eye of a single observer. In binaural synthesis, one acoustic signal is generated for each ear of a single listener and presented by headphones. These signals are designed to excite the eardrum as an acoustic sensor in the same way as a corresponding real sound field.

For the synthesis of this signal, a 3D model of the desired environment is used together with a specification of the acoustic properties of the desired sound source and the desired receiver in order to numerically simulate the sound propagation from source to receiver. The result of this simulation, including both the direct sound as well as all sound reflections at the boundaries of an open or enclosed space, is encoded as binaural room impulse responses (BRIRs), which describe the transfer path between a sound source and a defined (e.g.: human) receiver. If these impulse responses are linked to a temporal source signal such as speech or music by a numerical process called *convolution*, the corresponding sound source will be perceived at the desired point in space encoded in the BRIR. This perception includes all spatial properties of both the source and the environment, such as the distance, the direction of the source and the size and reverberation of the surrounding space. In order to invoke this perception, the result of the convolution has to be applied to the ear canal of a listener by headphones. The generated *auralization* (Vorländer 2008) of an acoustic scene will only be perceived as immersive if the listener can naturally interact

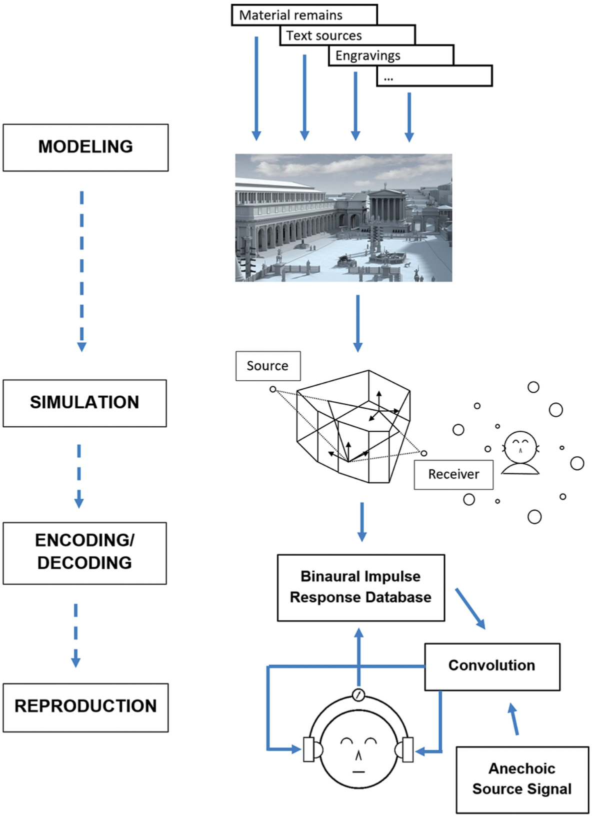

Fig. 2: Processing steps for the generation of virtual acoustic realities by binaural synthesis.

with the produced sound field, i.e. if the ear signals are re-calculated for different head orientations, so that the listener can move *within* the scene rather than the scene moving with the listener. This feature also has an equivalent in stereoscopic displays, in which the image has to be re-calculated in real-time whenever the observer changes his or her visual orientation. And just as with stereoscopic displays, it depends on the reproduction device, whether the simulation is presented as virtual or Augmented Reality. Headphones can exhibit different degrees of openness, ranging from closed headphones with a strong insulation against the external sound field, occasionally even supported by noise-cancelling algorithms, to open or even extra-aural headphones, which leave a gap of some centimeters between the ear and the headphone and constitute no significant obstacle for the external sound field. The latter are particularly suitable for simulations presented as Augmented Reality (see fig. 3).

Fig. 3: Violinist playing in an Augmented Reality with extra-aural headphones. The direct sound of his instrument, unobstructed by the open headphone design, is mixed with the simulation of a concert hall, generated by binaural synthesis.[2]

If the observations made within such a virtual acoustic reality shall be exploited as eyewitness reports about the corresponding historical scenery, the credibility of resulting inferences can be questioned at two levels. The first concerns the relationship between the simulation and the historical 'reality', the second concerns the reliability of the observations made by human observers inside the generated virtual or Augmented Reality. Both aspects will be considered more closely in the following.

2. Considering the uncertainty of virtual reconstructions

If virtual reconstructions are evaluated as a research tool, it is necessary to specify their epistemological function, since – as with any other tool – their adequacy

2 This setup was used to study how musicians adapt their way of playing to different room acoustic environments (cf. Schaerer Kalkandjiev/Weinzierl 2015).

can only be determined with respect to a desired goal. In the case of historical research, this requires a well-specified research question or hypothesis that can be answered of confirmed on the basis of qualitative or quantitative data retrieved from virtual historical environments. In the example discussed throughout this text, the question concerns the speech intelligibility in the ancient Forum Romanum, in other cases it might relate to the sound impression of historical concert halls, the visual impression of historical churches, the visual impression of paintings in museums before and after the introduction of electric light or other similar problems.

In the present discourse about the possibility of research in virtual environments, we think that the discussion is often overly focused on the question of whether simulations can replace the experience of a real, historical environment, or, more generally, whether virtual reality can fully replace 'real reality'. This question seems unproductive to the authors. This is less because it might be technically unfeasible in the foreseeable future to provide fully transparent virtual environments.[3] More relevant from a methodological point of view is that there is no criterion to answer this question without specifying the task, that an observer of (and listener 'in') the virtual or Augmented Reality has to perform. For most questions of mere perceptibility (e. g., speech intelligibility such in the present example), however, such criteria can be well-formulated and justified – based on theory or on empirical pretests – for a media system in order to fully convey the necessary cues for a listener confronted with this task.

Consequently, it might not always be necessary to generate virtual environments with a maximum level of realism and interactivity (whatever that implies). Simply with respect to efficiency and limited resources, it might be rather desirable to provide a virtual environment with only the features necessary to answer the question for which it was devised.

3. Reliability and bias I: The virtual environment

As far as the process of transforming historical evidence into the virtual reality presentation of a certain historical scenario is concerned, each step of the transformation can introduce uncertainties which affect the credibility of the final result. If we take a closer look at each of these steps in the order they have to be processed in practice (fig. 2), the first stage includes the establishment of an en-

3 The "perceptual illusion of non-mediation" was suggested as a criterion for completely transparent virtual realities providing an unrestricted sense of presence (Lombard/Ditton 1997, 9).

vironmental model of the desired scenario. In visual simulations this would include a 3D model incorporating the geometry of the space and the texturing of the surfaces. In acoustic simulations it would include a 3D model and the acoustic properties of all boundaries such as the absorption and scattering coefficients of the employed surfaces. The types of model parameters required at this stage depend on the subsequent simulation algorithm; while for wave-based numerical simulations of the sound propagation such as the finite element method (FEM) a complex impedance is required for each boundary element representing the amplitude and phase difference between the incoming and the outgoing sound wave, for ray-based acoustical simulations, often combining raytracing and image-source-method algorithms, only a real-valued reflection coefficient is required, as the phase of the acoustic wave is ignored in this simulation type.

No matter which simulation approach is selected, any of the required input parameters is subject to uncertainty which may result from insufficient knowledge of the historical conditions themselves or from insufficient knowledge of how to translate historical evidence into the required input parameters. In the case of the Forum Romanum, for example, there was insufficient knowledge about the material used for the walls of the surrounding buildings. Furthermore, also the absorption and scattering properties of Roman concrete (*opus caementicium*), which was introduced after 200 BC and used for many of the larger buildings of the forum, could only be estimated. The exact state of construction of the forum at a specific point in time was partly unknown, for there were contradictory information in different sources such as pictorial representations and textual sources. Further, while it was known that certain sound sources (speakers, horse-drawn vehicles, other sources of noise) were present in the investigated scenario, their exact acoustic radiation characteristics, i.e. the frequency-dependent directivity, might be unknown.

The uncertainties at this stage, which we suggest calling *modeling uncertainty*, result from the classical problem of incomplete or inconsistent historical evidence. They are not yet related to the computational processing of this evidence, but have to be understood and treated with a classical source-critical approach; the simulation method only defines the type of evidence required in order to generate a virtual representation. Historians might otherwise not be interested in the sound absorption of Roman concrete; for the acoustic simulation it is essential information.

At the second stage, the reliability of the numerical simulation itself may be questioned, i.e. the assumption that the sound propagation from the source to the receiver is correctly modeled by the selected numerical approach. We could term the uncertainties at this stage *simulation uncertainty.* Simulation uncertainty is an engineering problem inherent to any kind of numerical simulation of physical processes. It is usually addressed by using measurements conducted on real physical systems as a reference and by comparing these to simulations on the basis of specific dependent variables. These variables are selected according to a defined functional application. For example, if it is the function of an acoustic simulation to predict the reverberation of newly designed concert halls, measured reverberation times would constitute an appropriate reference for simulated reverberation times, as would other room acoustic parameters (Bork 2005a, 2005b). In the case of virtual realities intended to create a convincing sensory impression, the perceived difference to a corresponding real environment would be an appropriate reference for an evaluation of the simulation uncertainty. For the operationalization of the 'perceived difference', different measurable indicators have been proposed (see next section).

At the next stage, the results of a numerical simulation have to be encoded in a certain data format for storage or for live transmission, and decoded for reproduction. As with any communication channel, the information transmitted by encoding and decoding is overlaid by noise, which can have a multitude of sources. In the case of binaural synthesis, these include, for example, the spatial discretization of the sound field at the receiver. Whereas in a real sound field, listeners can introduce infinitesimally small modulations to the sound field by infinitesimally small head movements, BRIRs are available only for a predefined grid of head orientations. They have to be interpolated for head orientations in between. The simplest solution is a nearest-neighbor interpolation, i. e. the hard switching in the predefined grid; but even more advanced solutions are always error-prone approximations. Other encoding/decoding errors can be due to the fact that only initial parts of the impulse responses are exchanged dynamically to increase computational efficiency. They can also result from the audio signal format used or simply from the numerical resolution of the digital system used. The errors introduced by encoding and decoding could be summarized as *coding uncertainty*.

At a final stage, the accuracy of the optical or acoustic reproduction can be questioned, i. e. the extent to which the quality of the optical or acoustic signal presented to the eyes or the ears with respect to the desired reference is degraded by the employed reproduction system. Whereas coding uncertainty is introduced in the digital domain, for *reproduction* some kind of human-computer interface is required. In the case of binaural synthesis, degradations can relate to the headphones and the head tracking device used. The transfer function of the headphones, for example, can lead to timbral differences between the original and the technically reproduced sound source. Latencies in the acoustic adaption of the ear signals can lead to spatial instability of the reproduced sound source, latencies in the visual adaption in head mounted displays can even cause so-called simulator sickness if visual and vestibular motion cues are no longer consistent. Inaccuracies at this stage could be called *reproduction uncertainty*.

4. Characterizing the reliability of virtual environments

The uncertainties introduced at the modeling, simulation, coding and reproduction stage have to be indicated and quantified by appropriate means in order to convey a transparent overall picture of the reliability of the virtualization.

At the modeling stage, incomplete knowledge about the original spatial environment can, for example, be documented by alternative models representing different potential historical conditions, all of which can be plausible with respect to the archeological remains and historical knowledge in general. For modeling parameters such as the acoustic boundary conditions, an estimated range of possible values can be given according to expert knowledge. There are different ways to investigate the propagation of these uncertainties at the level of input parameters through the simulation to output parameters such as, in our example, speech intelligibility in the Forum Romanum. For discrete input parameters, such as different models representing different possible spatial configurations, a simulation can be run for every element of a set of models, giving a corresponding set of output parameters. For input parameters with a continuous estimated range of uncertainty, a sampling scheme may be applied, and then the simulation has to be run for this sample of input parameter values in order to estimate the output uncertainty. Finally, a statistically corroborated estimation of the uncertainties at the level of the output parameter would specify input parameters as probability distributions and use a sampling scheme that considers the probability of each parameter value. This can be achieved by random sampling (Monte Carlo strategies) or by dividing the range of each uncertain input parameter into equi-probable intervals and sampling accordingly (Rubinstein/Kroese 2008). The result is then a set of output values representing the probability distribution of the parameter of interest (cf. fig. 4).

The uncertainties arising from the numerical simulation algorithm itself are usually estimated by comparing simulations to measurements of the corresponding real physical systems. This is not as straightforward as it may seem, however, because the measurements used as references for the uncertainty of the simulation have a measurement uncertainty themselves. Moreover, it can be difficult to ensure that the input parameters used for the simulation will correspond exactly to the parameters of the real system. And, finally, benchmark tests between measurement and simulation are, at least in acoustics, based on considerable technical efforts. Hence, they can be conducted only for a small sample of physical systems and provide only a rough estimation of the reliability of the simulation algorithm without statistical information about the uncertainty distribution. Nevertheless, these benchmarks are common for most numerical simulation methods in the form of round-robin tests or competitions between alternative approaches and implementations.[4]

Encoding and decoding errors can be controlled by system design in order to keep them below a predefined technical or perceptual threshold. The spatial discretization of binaural simulations, for example, has been shown to be inaudible with a sampling grid of 2° for the horizontal, vertical and lateral discretization of BRIRs (Lindau/Weinzierl 2009). With respect to the dynamic exchange of these impulse responses for different head orientations, a prediction formula has been devised specifying the time window necessary for such a dynamic exchange in order for audible artifacts to be avoided (Lindau *et al.* 2012). The underlying test procedures are similar to those used for the encoding and decoding of audio signals with data compression (mp3), which determine a threshold of transparency, i. e. the data rate necessary for the differences between the encoded/decoded source material and a given reference to be inaudible or come with acceptable degradations.

At the reproduction stage, it is the established parameters for imaging devices or sound transducers that can be applied to evaluate the information loss between the computer model and the human user. These include the field of view or the image resolution of displays or the frequency bandwidth, the linearity of the transfer function or non-linear distortions caused by headphones. For interactive virtual environments, it includes the update rate or the system latency of the head tracking device. For all these parameters there is a large stock of research in the field of quality and usability research that can be used in order to assess the perceptual relevance of the intrinsic limitations of the systems and devices used.

For virtual acoustic environments as a whole, different degrees of distinctness of simulation and reality have been suggested as measures: The *plausibility* of a simulation (Lindau/Weinzierl 2012), measuring the ability of a subject to identify the simulation in a random, alternating presentation of simulation and reality, and the *authenticity* of a simulation (Brinkmann *et al.* 2014), measuring the ability of a subject to perceive *any* difference between simulation

4 For examples see http://rr.auralisation.net/ for room acoustic simulation and auralization, and Durante/Riedel (2008) for wind flow simulation.

and reality, even if the simulation cannot be identified as such. For a differential diagnosis of differences between simulation and real references, a taxonomy has been developed by qualitative research (Lindau *et al.* 2014). While these tests give no indication of the stage at which the virtualization degradations occurred, they give an overall picture of the quality provided.

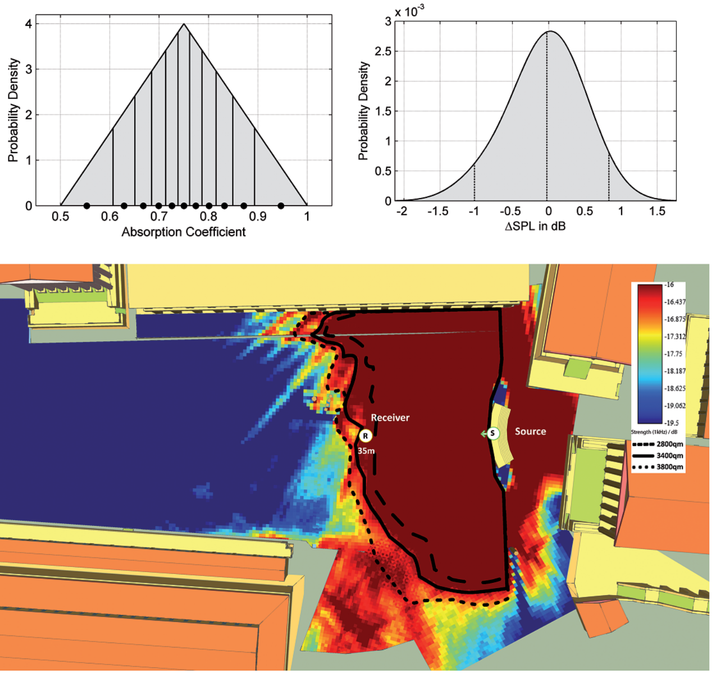

Fig. 4: Uncertainty propagation through the acoustic simulation. Top: Assuming that the absorption coefficient of the audience lies in an interval between 0.5 and 1.0 (with a triangular distribution, and 0.75 as the most probable value, top left), the resulting uncertainty of the sound power level at the border of intelligibility is ca. ±1 dB (5 %/95 % quantiles, indicated as dashed lines, top right). Bottom: The color map shows the regions of good (red) and poor (blue) speech intelligibility on the ancient Forum Romanum. The solid line indicates the border of intelligibility, with the dashed lines related to the uncertainty of ±1 dB sound power level calculated above. The corresponding size of the audience area that could be reached by a trained speaker is 3400 m^2 (2800 m^2/3800 m^2).

5. Reconstructing historical events vs. reconstructing historical *spaces of possibility*

The above mentioned uncertainties at the different stages of digital reconstructions will, as a result, lead to a corresponding range of uncertainty for the specified variables of interest, such as speech intelligibility in the described model study. Independent of this problem is the question of whether the generated virtual environment and the results derived from it are supposed to represent one specific historical scenario, or whether they aim at a *range of possibilities* related to a certain historical situation.

Are we interested in the size of the audience able to understand Marcus Tullius Cicero during his speech *In Catilinam* on November 8, 63 BC, or are we interested in how many persons could *typically* understand speeches given by different orators on different occasions during a given historical period? For the first problem, there will usually be a high degree of modeling uncertainty, because we do not know exactly how powerful the voice of this specific person was, what the rhetorical concept of his speech was, or how large and how attentive the crowd was on this specific occasion. Hence, we will be confronted with a high modeling uncertainty leading to a corresponding uncertainty about the level of the results on speech intelligibility. For the second problem, the historical scenario itself is described by a range of input parameters rather than a single model state, and the limits of this range can usually be determined more precisely than the conditions of one specific event. We can determine quite precisely the range in which the sound power of trained speakers lies (e. g. through empirical investigations on present-day subjects), by assuming that the size of the crowd could be something between virtually zero and an audience corresponding to the maximum density of standing persons possible in a public place, or by taking into account the fact that the level of attention of an audience at public speeches in terms of the background noise level can again be empirically determined, assuming that the noise level produced by a modern crowd is not substantially different from a crowd during the Roman period.

In order to describe a range of historical conditions corresponding to a state space of computer models instead of one specific historical occasion, similar strategies to those being used to describe a range of modeling uncertainties can be applied, i. e. by calculating the model for the limits of the range or by estimating a probability distribution for the respective input parameter and by letting a sampled version of this distribution propagate through the simulation in order to calculate the resulting uncertainty at the level of the parameter of interest (fig. 4). In practice, due to the limited information about single events and the empirically more substantiated knowledge about the range of historical conditions, there will often be a trade-off between modeling uncertainty and historical specificity. For the results that virtual reconstructions may provide for a defined range of historical events are usually much more reliable and thus scientifically more valuable than those for specific historical events.

6. Reliability and bias II: The observer

Unlike a merely numerical simulation, a virtual historical environment (VHE) encompasses interactive real-time sense data (sometimes also autonomous artificial agents) that are experienced and interpreted by human agents in order to make new knowledge claims about the historical circumstances under simulation. To legitimate such claims, even in the face of a technically perfect simulation of historical reality, scientists employing such a strategy should always follow established scientific quality criteria for conclusions drawn from systematic observations. After all, human agents tend to be unreliable, unique and autopoietic measurement instruments. They have constrained attention as well as different cognitive-perceptual capacities and they construct the identity of objects, events and their meaning from the background of their historical cultural upbringing and embeddedness. Furthermore, in interactive environments, different human agents may by definition perform different actions, thereby also arriving at idiosyncratic impressions of the very same simulated worlds. These epistemological problems are well known in research disciplines that have a tradition of dealing with subjective field observations, such as psychology, sociology and ethnography. Within these disciplines, several methodological answers have been developed to address these challenges. Even the problem of systematically performing social science in simulated virtual realities has been discussed for nearly 20 years now (Markham 1998). However, the type of systematic observation performed within VHEs differs radically from psychological v*irtual reality experiments* (Veiling *et al.* 2013) as well as from c*yber ethnography* (Hallett/Barber 2014) in sociology or communication studies in that it is not directed at analyzing the actions or impressions of *other* human agents within virtual worlds or computer games but at analyzing a scholar's *own* subjective experience of being an actor within a simulated past cultural setting. Therefore, the methodological approach discussed here may be considered a new form of empirical inquiry that could be termed *cyber phenomenology.*[5] Nevertheless, we suggest building on

5 See Houliez/Gamble (2012) for a brief introduction into the phenomenological approach.

methodological techniques developed within the abovementioned fields, since, as we will try to show in the following, they help to improve and demonstrate the reliability and intersubjective validity of systematic scientific observations performed within VHEs.

7. Performing (semi-)structured phenomenological observations in VHEs to increase reliability

The most important challenge for cyber phenomenology from a *reliability perspective* appears to be the selectivity of human perception, regardless of cultural background and expertise. *Attention span*, cognitive *priming and halo effects* (Kochinka 2010), as well as the results of different actions performed in the same virtual environment by different actors may lead to very different subjective percepts. A straightforward methodological approach that is able to 'tame' the resulting idiosyncrasy of impressions with a long history of success in the social sciences is performing structured or semi-structured observations (Bryman 2008b). Again, these require a well-specified research question and a clearly specified range of phenomena of interest at the outset, which are then used to create an observation form (or category sheet) that formulates different activities a researcher has to perform (e. g. visiting certain places within the VHE) and leaves empty slots in the table to be filled out during or after the experience. These could be used to describe the subjectively experienced *mere existence* or the experienced *qualities or intensities* of certain a-priori expected phenomena. In the fully structured variant, all phenomena and their subcategories are already specified at the outset of the fieldwork, turning the table form into a multiple-choice questionnaire. In the semi-structured version, the categories for phenomena to look out for tend to be rather broad and the subcategories are open-ended, which leaves their creation to the field-observer, and thereby introduces an interpretive-integrative step in the later analysis (Gehrau 2002). Regardless of which of the two approaches has been adopted, due to their degree of standardization, both procedures principally allow the observations to be converted into numerical variables in a later step and thereby enable a quantitative comparison between the impressions of a larger sample of different historical researchers. This paves the way for systematic mathematical reliability checks, as well as for performing scale building and statistical hypothesis tests (Bryman 2008b). While internal validity and reliability are improved considerably when performing such structured observations, a drawback of these procedures is that they are only able to deal with a-priori expected phenomena, thereby imposing a deductive inferential style of inquiry with the typical tradeoff in external validity. Taken together, structured observation techniques appear well suited for quasi-experimental historical studies using VHEs that try to test specific hypotheses.

8. Employing open-ended ethnographic techniques in VHEs to increase intersubjective validity

The most important challenge for cyber phenomenology from a *validity perspective* appears to be the cultural-historical situatedness of human perception: every observational act is laden with implicit culturally-specific assumptions that structure the kind of entities, qualities, actors, relationships and events constructed by the human mind during perception (Bogen 2014). This problem even increases when research questions are of an exploratory nature. Therefore, from the standpoint of validity it appears important in these cases to employ a form of inquiry that maximizes the intersubjective traceability of subjective interpretations in order to 'tame' perception's cultural boundedness. Obviously, this is only possible if not the experienced phenomena alone, but also their subjective interpretations and the way they were reached are analyzed and compared qualitatively between different researchers with different backgrounds and areas of expertise. However, this is not possible with standardized observation forms (Bryman 2008a). Hence, when dealing with more open-ended research questions or when expecting a possible cultural-historical bias on the part of the scientific observer, it seems reasonable to accept drawbacks in reliability and to try to improve intersubjective validity by drawing on more 'qualitative' ethnographic techniques developed in cultural studies and ethnography instead of performing structured observations (Hallett/Barber 2014). Typical examples would be historical research projects that are not only interested in questions of the mere perceptibility or intensity of phenomena, but also in certain forms of aesthetic, emotional or social judgments. Hence, to ensure the intersubjectivity of the phenomenological findings from these scenarios, we recommend that researchers try to formulate "thick descriptions" (Geertz 1973) of their subjective impressions in the form of extensive written field notes or by employing think-aloud techniques (Bryman 2008a). To increase traceability of this subjective interpretive data, it should be enriched by locative data (positions, distances and actual movement in the VHE) and audiovisual-field recordings (screenshots, short movies and audio renderings) which would help other scholars to understand and trace the personal interpretations and conclusions. Furthermore, researchers performing cyber phenomenology in this way should explicitly disclose the stocks of historical cultural knowledge that they actively employ when trying to put themselves into the hermeneutic position of a contemporary observer. By triangulating these different forms of data in an interpretive analysis, which may also compare the field notes of a smaller number of researchers visiting the same scenario, a *grounded theory* (Bryant/Charmaz 2011) about previously unknown mechanisms and phenomena in the field may then be developed. Taken together, qualitative ethnographic techniques appear well suited for more open-ended, explorative historical studies using VHEs that try to arrive at new hypotheses.

9. Combining structured observations and ethnographic techniques

Obviously, ethnographic techniques and structured observations can also be combined in a complementary way to accommodate for each of their tradeoffs. Taking the research example discussed above, a historical scientist could have first ethnographically explored several VHEs with different historical versions of the Forum Romanum to arrive subsequently at the grounded theory that some of the positional changes of the Rostra must have been related to acoustic problems. This could give rise to an additional study employing structured observations in the same VHE in order to test this hypothesis statistically.

10. New knowledge based on existing evidence?

A final remark will be made on the possible future role of virtual environments and the related observation techniques we envision within an epistemological model of historiographical research. If the empirical material collected by historians is regarded as material traces of historical events that give rise to collective inferences regarding ‘common causes’ (Tucker 2004), one still has to explain the additional value of virtual environments, when these can only be based on evidence that is already known. Every aspect of the example simulation of the virtual Forum Romanum discussed in this article, from the forum’s architecture and the surrounding buildings to the location of the Rostra was based on previous knowledge in classical archeology, and the computer model did not bring to light any new historical evidence.

Nevertheless, we see three ways in which new knowledge could emerge from virtual or augmented historical environments. First, by systematically relating together scattered bits of information (about the architecture, the architectural materials used, the location of the Rostra, the text of the speech) within a model of sound or light propagation, new information can be derived that was already embedded in the historical evidence available but had not been combined before. Thus, the scientific theories about sound or light propagation take on the role of *nomic regularities*, which carry "nested information" about the source of information, i.e. the original historical events (Tucker 2004, 18 f.).[6] This is similar to the application of natural laws about the ageing of materials in order to extrapolate the original appearance of the materials from the existing remains together with knowledge about the climatic conditions of the past. Since disciplines such as geology, meteorology, the material sciences or acoustics are currently intensively involved with the development of computer models for the dynamic, i.e. time-related, behavior of physical systems, the historical sciences would do well to exploit the wealth of nomic regularities inherent in these models.

Second, since virtual or augmented environments are able to lift the numerical variables of computer models to the level of sensory signals, they not only convey more comprehensive information on questions concerning the perceptual impression of historical environments; these are also easier to access explanatorily for non-experts in the respective scientific domain. A prediction of the speech intelligibility in the ancient Forum Romanum could, in principle, already be made based on information about the sound power of trained speakers, the distance between speaker and audience, the reverberation time of the place and the noise level generated by the audience. However, the analytical relations involved are themselves derived from psychophysical experiments. As statistical correlations, they have only limited accuracy when predicting the intelligibility of presentations of texts in a specific language with a specific rhetorical strategy from several listening perspectives. And it is one thing to know that the predicted speech transmission index (STI) at a certain place in the audience has a value of 0.7 and another thing to be able to listen to the virtualization of the speech in different places in the audience – at least for listeners who are not familiar with the meaning of STI values. Insofar, VHEs help to explore the actual perceptual meaning of coefficients derived from mere numerical simulations.

6 At this point, Tucker applies Dretske's information theoretical concept of knowledge and knowledge production to the historical sciences (cf. Dretske 1981, 71 f).

And finally, the integration of knowledge stocks from different academic disciplines, required for the empirically substantiated design of virtual historic environments, might make these a focal point of interdisciplinary cooperation, where scientific knowledge from areas as far apart as classical archeology, architectural and building history, the history of rhetoric, room acoustics and speech acoustics converge at one common point. And the history of science is full of examples of new knowledge emerging from a new and unusual cooperation between different disciplines.

References

Bearman, D. (2011): "3D Representations in Museums." In: *Curator: The Museum Journal* 54 (1), 55 – 61.
Bogdanovych, A./Rodriguez-Aguilar, J. A./Simoff, S./Cohen, A. (2010): "Authentic Interactive Reenactment of Cultural Heritage with 3D Virtual Worlds and Artificial Intelligence." In: *Applied Artificial Intelligence* 24(6), 617 – 647.
Bogen, J. (2014): "Theory and Observation in Science." In: E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2014). http://plato.stanford.edu/archives/sum2014/entries/science-theory-observation/ (accessed 10. 1. 2016).
Bork, I. (2005a): "Report on the 3rd Round Robin on Room Acoustical Computer Simulation, Part I: Measurements." In: *Acta Acustica united with Acustica* 91(4), 740 – 752.
Bork, I. (2005b): "Report on the 3rd Round Robin on Room Acoustical Computer Simulation, Part II: Calculations." In: *Acta Acustica united with Acustica* 91(4), 753 – 763.
Brinkmann, F./Lindau, A./Vrhovnik, M./Weinzierl, S. (2014). "Assessing the Authenticity of Individual Dynamic Binaural Synthesis." In: *Proc. of the EAA Joint Symposium on Auralization and Ambisonics, Berlin*, 62 – 68.
Bryant, A./Charmaz, K. (2011): "Grounded Theory." In: M. Williams/W. P. Vogt (ed.), *The SAGE Handbook of Innovation in Social Research METHODS*, Los Angeles et al.: SAGE, 205 – 227.
Bryman, A. (2008a): "Ethnography and Participant Observation." In: *Social Research Methods. 3rd Edition*, Oxford: Oxford University Press, 400 – 434.
Bryman, A. (2008b): "Structured Observation." In: *Social Research Methods. 3rd Edition*. Oxford: Oxford University Press, 253 – 272.
Dretske, F. I. (1981): *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.
Durante, F./Riedel, V. (2008): *Round Robin Numerical Flow Simulation in Wind Energy. Final Report.* http://www.dewi.de/dewi/fileadmin/pdf/publications/Studies/Round%20Robin/round-robin-flowsimulation-final-report.pdf (accessed 10. 1. 2016).
Geertz, C. (1973): "Thick Description: Toward an Interpretive Theory of Culture." In: *The Interpretation of Cultures: Selected Essays*, New York: Basic Books, 3 – 30.
Gehrau, V. (2002): *Die Beobachtung in der Kommunikationswissenschaft: Methodische Ansätze und Beispielstudien*, Konstanz: UTB GmbH.
Gernsheim, H./Gernsheim, A. (1969): *L.J.M. Daguerre: The History of the Diorama and the Daguerreotype*, Dover: Dover Publications.

Hallett, R. E./Barber, K. (2014): “Ethnographic Research in a Cyber Era.” In: *Journal of Contemporary Ethnography* 43(3), 306 – 330. DOI: 10.1177/0891241613497749.

Houliez, C./Gamble, E. (2012): “Dwelling in Second Life? A phenomenological evaluation of online virtual worlds.” In: *Virtual Reality* 17(4), 263 – 278, DOI: 10.1007/s10055-012-0218-1.

Kochinka, A. (2010): “Beobachtung.” In: G. Mey/K. Mruck (eds.), *Handbuch Qualitative Forschung in der Psychologie*, Berlin: Springer, 449 – 461, http://link.springer.com/chapter/10.1007/978-3-531-92052-8_32 (accessed 10. 1. 2016).

Lercari, N. (2016): “Simulating History in Virtual Worlds.” In: Y. Sivan (ed.), *Handbook on 3D3C Platforms*, Berlin: Springer, 337 – 352.

Lindau, A./Erbes, V./Lepa, S./Maempel, H.-J./Brinkmann, F./Weinzierl, S. (2014): “A Spatial Audio Quality Inventory for Virtual Acoustic Environments (SAQI).” In: *Acta Acustica United with Acustica* 100, 984 – 994.

Lindau, A./Weinzierl, S. (2012): “Assessing the Plausibility of Virtual Acoustic Environments.” In: *Acta Acustica United with Acustica* 98, 804 – 810.

Lindau, A./Kosanke, L./Weinzierl, S. (2012): “Perceptual Evaluation of Model- and Signal-based Predictors of the Mixing time in Binaural Room Impulse Responses.” In: *Journal of the Audio Engineering Society* 60(11), 887 – 898.

Lindau, A./Weinzierl, S. (2009): “On the Spatial Resolution of Virtual Acoustic Environments for Head Movements in Horizontal, Vertical and Lateral Direction.” In: *EAA Symposium on Auralization*, Helsinki, 1 – 6.

Lombard, M./Ditton, T. (1997): “At the Heart of It All: The Concept of Presence.” In: *Journal of Computer-Mediated Communication* 3(2). DOI: 10.1111/j.1083 – 6101.1997.tb00072.x.

Lombardo, V./Valle, A./Fitch, J./Tazelaar, K./Weinzierl, S./Borczyk, W. (2009): “A Virtual-Reality Reconstruction of Poéme Électronique Based on Philological Research.” In: *Computer Music Journal* 33(2), 24 – 47.

Markham, A. N. (1998): *Life Online: Researching Real Experience in Virtual Space*, Lanham: Rowman Altamira.

Mikropoulos, T. A./Natsis, A. (2011): “Educational virtual environments: A ten-year review of empirical research (1999 – 2009).” In: *Computers & Education* 56(3), 769 – 780.

Oettermann, S. (1980): *Das Panorama. Die Geschichte eines Massenmediums,* Frankfurt on Main: Syndikat.

Pujol, L./Champion, E. (2012): “Evaluating Presence in Cultural Heritage Projects.” In: *International Journal of Heritage Studies* 18(1), 83 – 102.

Rubinstein, R. Y./Kroese, D. P. (2008): *Simulation and the Monte Carlo Method*, New York: Wiley.

Schaerer Kalkandjiev, Z./Weinzierl, S. (2015): “The Influence of Room Acoustics on Solo Music Performance: An Experimental Study.” In: *Psychomusicology: Music, Mind & Brain* 25(3), 195 – 207.

Stogner, M. B. (2009): “The Media-enhanced Museum Experience: Debating the use of Media Technology in Cultural Exhibitions.” In: *Curator: The Museum Journal* 52(4), 385 – 397.

Tucker, A. (2004): *Our Knowledge of the Past. A Philosophy of Historiography*, Cambridge: Cambridge University Press.

Veling, W./Brinkman, W.-P./Dorrestijn, E./van der Gaag, M. (2013): “Virtual Reality Experiments Linking Social Environment and Psychosis: A Pilot Study.” In: *Cyberpsychology, Behavior, and Social Networking* 17(3), 191 – 195. DOI: 10.1089/cyber.2012.0497.

Vorländer, M. (2008): *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, Berlin: Springer.