

**C. Barot\***  
**D. Lourdeaux**  
Heudiasyc UMR CNRS 7253  
Université de Technologie  
de Compiègne  
BP 20529  
60205 Compiègne, France

**J.-M. Burkhardt**  
IFSTTAR LPC  
Versailles, France

**K. Amokrane**  
**D. Lenne**  
Heudiasyc UMR CNRS 7253  
Université de Technologie  
de Compiègne  
BP 20529  
60205 Compiègne, France

# V3S: A Virtual Environment for Risk-Management Training Based on Human-Activity Models

## Abstract

As work situations become more complex, virtual reality has proven to be advantageous for the emerging needs in risk-management training. However, building perfectly realistic simulations of the technical systems is not enough to address complex coactivity situations, where human factors play an important part. There is a need for virtual environments that would put trainees into varied ecological situations, inducing knowledge and competencies that would be put into practice in genuine work situations. The V3S project was proposed to build a generic framework for tailor-made virtual environments that can adapt to different application cases, technological configurations, or pedagogical strategies. This framework relies on the integration of multiple explicit models (domain, activity, and risk model). In order to build ecologically valid virtual environments, these models represent not only the prescribed activity, but the situated knowledge of operators about their tasks, including deviations from the procedures. These models are used both to monitor learners' actions, detecting errors and compromises; and also to generate virtual characters' behaviors, subject to erroneous actions. Moreover, dynamic situated feedback allows for progressive learning scenarios, adapting the complexity of the situations to the learner's activity and level. Evaluations have shown a high satisfaction level and encouraging usability measures. In future work, we propose to extend the possibilities of the simulation through the creation and monitoring of adaptive scenarios, adjusting the behavior of virtual characters able to assist or disrupt the user.<sup>1</sup>

## 1 Introduction

In high-risk industries, risk-management training has become a major issue due to several factors, and virtual reality has been shown advantageous for addressing this problem (Fabre, Couix, Burkhardt, Gounelle, & Cabon, 2006) in an ecological<sup>2</sup> approach. Rather than aiming for perfectly realistic

\*Correspondence to [camille.barot@hds.utc.fr](mailto:camille.barot@hds.utc.fr).

1. This paper is an extended version of *V3S, A Virtual Environment for Risk Management Training*, by C. Barot, J.-M. Burkhardt, D. Lourdeaux, and D. Lenne, the recipient of the Best Paper award at the Joint Virtual Reality Conference of EuroVR, EGVE (2011).

2. Initially introduced in the domain of psychology of perception, the notion of ecological validity refers to experimental situations that approximate the real-life situation under study. It has been contrasted with controlled and contrived situations in sterile and artificial situations often referred to as "laboratories" (see Loomis, Blascovich, & Beall, 1999; Burkhardt, Bardy, & Lourdeaux, 2003, for an extensive discussion of the components of ecological validity in experiments using virtual environments).

environments, this approach chooses objects and actions, as well as their representations within the virtual environment, in order to foster the learning of knowledge structures similar to those required in practice in genuine work situations.

However, existing virtual environments focus on precise tasks, such as the repetition of a technical gesture or a prescribed procedure, in situations with little complexity. Yet, when addressing complex, nonideal situations, where human factors are predominant, it is no longer possible to predict all possible sequences of events; thus, existing pedagogical control systems can no longer be used.

Our hypothesis is that activity models of humans at work can provide a sound basis to specify learning situations, paths, and goals. Although this approach is not new from a psychological perspective (Annett & Duncan, 1967), so far it has only been applied to a handful of virtual environments for learning and training (Burkhardt et al., 2003; Leva et al., 2009). The main reason is the lack of a framework and tools to support the actual use of such models in virtual environments. Yet, in such a framework, these models could be used and adapted regarding different pedagogical strategies as well as in the context of different interactive technologies. Having explicit instead of implicit models would allow experts from different fields to work in parallel and iteratively on these models. Moreover, the same model could serve as a basis for numerous different modules.

In the V3S (Virtual Reality for Safe Seveso Sub-contractors) project, this approach has been used to design a virtual environment, taking human factors into account. It consists of a generic framework which can be tailored to fit different training needs through the use of ergonomic and risk analyses. The application is equipped with a learner monitoring module capable of controlling the events happening in the environment according to pedagogical rules. It is also populated with virtual autonomous characters subject to mistakes and compromises.

Section 2 gives a brief overview of risk-management training context, the use of virtual reality for training and risk management, and the difficulties they meet to adapt to different pedagogical strategies. Section 3

covers the different components of the V3S project and Section 4 covers the evaluation study. Finally, Section 5 is a discussion of the work, including enhancing the system's adaptability through the introduction of dynamically adaptive scenarios in the simulation. Section 6 concludes the work.

## 2 Context and Previous Work

### 2.1 Issues in Risk-Management Training

The convergence of an increasing complexity of industrial processes, and having fewer and fewer operators with long-term experience, makes it challenging to manage actions and decision-making in abnormal situations. The precariousness of work and operator turnover have also produced loss of expertise. Finally, the design of human-machine interfaces has been shown to affect the capability of operators in industry to develop and maintain accurate knowledge about the systems and processes they are working on (Bainbridge, 1987; Jamieson & Vicente, 2001) and even to generate some specific risks due to negative transfers between variations of the systems (Besnard & Cacitti, 2005). Thus, it may occur that teams become incompletely aware of the phenomena underlying the processes they are monitoring. Training faces the challenge of providing the different intervening operators with sufficient competence in reasonable time periods for very complex systems.

Subsequently, training should in particular be oriented toward (1) how to cope with emergency/abnormal situations, (2) improving the understanding of the process, (3) initial training and skill sustenance on dangerous equipment and rare situations, and (4) dissemination of good/best practices, rules, and procedures. Training also closely complements other safety approaches, such as safety management (planning, organizing, controlling, monitoring, auditing, and reviewing the intervention process to reduce risk in a plant or in risky working conditions), risk assessment a priori, and a posteriori accident investigation (Fabre et al., 2006).

Obviously, no tool, not even virtual reality (VR), could single-handedly solve all these problems. Virtual

environments (VEs) have to be designed as part of a set of tools and pedagogical approaches, and the role of the trainer in this design process is crucial in order for the VE to be both efficient and adapted for use in training sessions. This implies considering both the domain and the type of pedagogical situation that the tool addresses.

## 2.2 Virtual Reality for Training and Risk Management

For a long time, operator training has been shown as the main expected use of VR in industries dealing with complex systems and the management of critical risk and safety issues (Fabre et al., 2006). Several orientations and learning objectives can be distinguished when using VR for training (Fabre et al.; Gounelle et al., 2007):

1. Training for emergency situations: several domains have been previously addressed in the literature, such as medical emergency (Stytz, Banks, Garcia, & Godsell-Stytz, 1997), military peacekeeping (Swartout et al., 2006), gas emission in refinery operations (Haller, Kurka, Volkert, & Wagner, 1999), Seveso site accidents (Edward, Lourdeaux, Lenne, Barthes, & Burkhardt, 2008), or terrorism (Li, Zhang, Xu, & Liu, 2006).
2. Improving the understanding of the dangerous/critical process or situation to work with, as, for example, in chemical reaction engineering (Bell & Fogler, 2004).
3. Providing initial training (Johnson, 1998; Ustarroz et al., 2004) and skill maintenance on dangerous equipment and rare situations.
4. Training in safety and hazard detection/evaluation: examples involve the training of health and safety prevention engineers (Gardeux, Marsot, & Rolin, 2005), students in engineering enrolled in a senior plant design course (Bell & Fogler, 2000), training of traffic accident investigators in the procedures to apply when intervening in an accident (BinSubaih, Maddock, & Romano, 2006), and so on.

In Burkhardt, Lourdeaux, and Lequatre (2005), we have shown based on a set of 17 existing VEs that VEs

for training can be categorized according to their provided functions, interaction devices, user representation, user feedback, offered guidance, and their intended pedagogical uses and situations. There are, more precisely, four main emerging patterns regarding pedagogy and the interactional configuration of learner(s) and trainer(s):

1. Tutoring-oriented VE, where one trainee is interacting with the environment, and one or more tutors are monitoring the session and providing guidance and didactic interaction during the training session; this pattern has been found for the most part to support gestural, behavioral, and procedural learning. There is no learning-specific guidance or feedback, partially because the design rationale is to mimic the feedback occurring in the real situation. This type of system uses various interaction devices ranging from desktop to fully immersive solutions with visual, sound, and haptic displays.
2. VE to support supervision of group-based activities by trainers, where learners perform activities in a group under the monitoring of one trainer, either at a distance or in colocation. This pattern has been mostly found in two training domains: command and crisis management, and school education. This pattern is based on simple interaction devices (such as mice or joysticks) dedicated to activities in a group, often in association with large screens to share views and to support collaboration.
3. VE as a rich media to support teaching in front of a class/audience, where control is essentially handled by the trainer, even if it can be given to one trainee at some point in the lesson. This pattern is mostly encountered to support lectures in highly visualization-oriented domains, as well as training on diagnosis in complex systems. Interaction devices are as simple as possible (3D mouse, PC mouse, joystick) and are used with a large common visual display.
4. VE for self-training, where the focus is on providing the learners with a whole environment to support individual autonomous learning.

Although these four technological configurations have been observed in current safety management and training-oriented applications (e.g., desktop VR, immersive room), most of them are desktop-low-cost VR systems, although an increasing number of immersive multi-user configurations are being developed.

It is important to make clear the nature of models that are actually integrated and used to animate the session of interaction provided to learners and trainers. Adopting a human-factors (HF) perspective, there can be three main classes of models integrated within a safety-oriented VE (Fabre et al., 2006):

1. Models that represent physical and technical components in the simulated situation and environment. The scope of this class covers processes and/or physical phenomena. This may be a simulation of plant processes, plant layouts, or chemical reactions, among other options. Examples of VE functions involving this class of models are to provide learners with a process simulation, to support the visualization and navigation within a plant to detect specific hazards, to enable operators to train on operating step-by-step on a simulated machine, and so on.
2. Models that explicitly refer to human-factors models and/or guidelines, for example, anthropometrics, cognition biases, team cooperation, task/activity of virtual operators, and organizational models. Examples of VE applications requiring this class of models are providing designers with HF guideline assistance during the design phase of a plant, or providing trainees with simulated virtual operators they should collaborate with in order to achieve their task or to cope with an emergency case study.
3. Models that enable the representation or labeling of specific safety objects called barriers (Hollnagel, 1994). Barriers are obstacles, obstructions, or hindrances that may prevent an action from being carried out or an event from taking place, as well as prevent or lessen the impact of the consequences. Several types of barriers can be modeled; for example, physical barriers (e.g., a protective wall, safety

glasses to avoid projections), functional barriers; that is, barriers which require preconditions to be satisfied (e.g., a lock requires a key), symbolic barriers (e.g., traffic signs, warnings), and finally, incorporeal (or organizational) barriers; that is, barriers which are not physically present but are on operators' minds (e.g., procedures, rules). Examples of VE applications could be to train safety analysts by providing a means to test and evaluate a plant subsystem with real or simulated operators after having introduced (or removed) a barrier (new descriptives, new rules or new procedures, etc.), or to train team managers and subcontractors on intervention planning by assessing the effect of adding or removing barriers in terms of safety consequences.

### 2.3 Pedagogical Scenario Control in Virtual Environments

Most virtual environments designed for training purposes are used in training sessions as part of a pedagogical scenario. These scenarios are the sequences of learning activities that the user has to perform. In some cases, several learning activities are simulated inside the virtual environment. The pedagogical scenario in the virtual environment generally consists of branching tree structures, containing predetermined sequences of scenes (Magerko, Wray, Holt, & Stensrud, 2005), or tasks that the user has to execute (i.e., the prescribed procedure; Mollet & Arnaldi, 2006). Yet, in order to stay within these paths, these virtual environments offer strong guidance to the trainee, often stopping the training whenever a mistake is made or when an action is taken that does not belong to the training scenario. To enable several pedagogical strategies, such as a trial-and-error approach, the system has to be able to cope with user agency.

In other cases, some environments are used in pedagogical scenarios as a single learning activity. These environments opt for the sandbox approach, letting the user act freely as the simulation evolves and reacts to each action. In these environments, such as in Shawver (1997), the only pedagogical control is that of the ini-

tial state of the world. However, without any real-time pedagogical control, the efficiency of the training is not guaranteed. The simulation could go in any direction, yet we would want it to be relevant to the profile and the current state of the trainee; without control, it is impossible to ensure a gradual learning process, whether it concerns the knowledge and competencies that the user has to demonstrate, the complexity of the situation, or the gravity of the consequences of his or her errors. For novice learners, for instance, it could be relevant to stop chain reactions before the situation gets out of hand.

One approach for ensuring both user agency and pedagogical control is to define a multilinear graph of all possible scenarios, adapting the reactions of the world and the virtual characters to the choices made by the user. In Delmas, Champagnat, and Augeraud (2007), the set of possible plots is thus explicitly modeled through a Petri network. However, when the complexity of the work situation scales up, it becomes difficult to predict all possible courses of action. Especially when the training addresses colocated activity situations, the computational complexity resulting from characters' interactions can be high. Yet the decision-making processes and emotions expressed by these characters have to stay consistent, even more so when they are aimed at reflecting ecological behaviors and are based on complex psychological models. In this case, it becomes impossible to foresee all possible combinations, and the virtual characters have to be given some autonomy for the scenarios to emerge from their actions.

Combining pedagogical control with user agency, all the while ensuring that the simulation stays coherent with regard to the models it is based on, is fundamentally problematic when complex situations are addressed. Rather than making explicit the whole set of possibles, the approach that has been adopted for the control of the training sessions in the V3S project involves the dynamic computation of situated feedback that varies depending on pedagogical rules. In particular, we offer situated scenario adaptation, modifying the technical system's reactions to create situations of an appropriate complexity given learning objectives and learner traces.

### 3 The V3S Project

The V3S project addresses the simulation of dangerous work situations related to maintenance activities performed by external companies on high-risk Seveso sites. Because of the subcontracting, these situations are prone to particular problems. Indeed, the manager's lack of control of the subcontractor's activity leads to a very formal application of procedures, which does not guarantee true risk control. Thus, industries are in need of tools to promote reflection from the different actors (decision-makers, managers, operators, etc.) in order to induce a true understanding of the risks that are at stake.

The V3S framework consists of a set of generic modules (see Figure 1). It can be adapted to different applications by changing the models used as the input and by adding or removing modules, in order to create tailor-made virtual environments that would be adapted to learning objectives and the needs of the learners and trainers. In order to enhance the ecological validity of the resulting simulations, several models of human activity in risk situations are considered. In particular, these models integrate the consideration of errors in human activity, whether they come from the learner or from the virtual characters who populate the environment.

#### 3.1 Models

Most virtual environments solely take into account models of simple technical systems, leaving aside activity or risk modeling (Fabre et al., 2006). Indeed, computing complex models created by domain experts, activity analysts, or ergonomists is problematic: either the computer scientists creating the simulation have to enter these models themselves, and thus have to think about how the system works; or else they have to go through a tedious pass of translating and unifying the different models provided by the experts toward a single computable formalism. Indeed, even if these models concern different views on the work situation, they include chunks of knowledge (such as objects, events, etc.) that computer modules have to be able to link together in order to reason on more than one model.

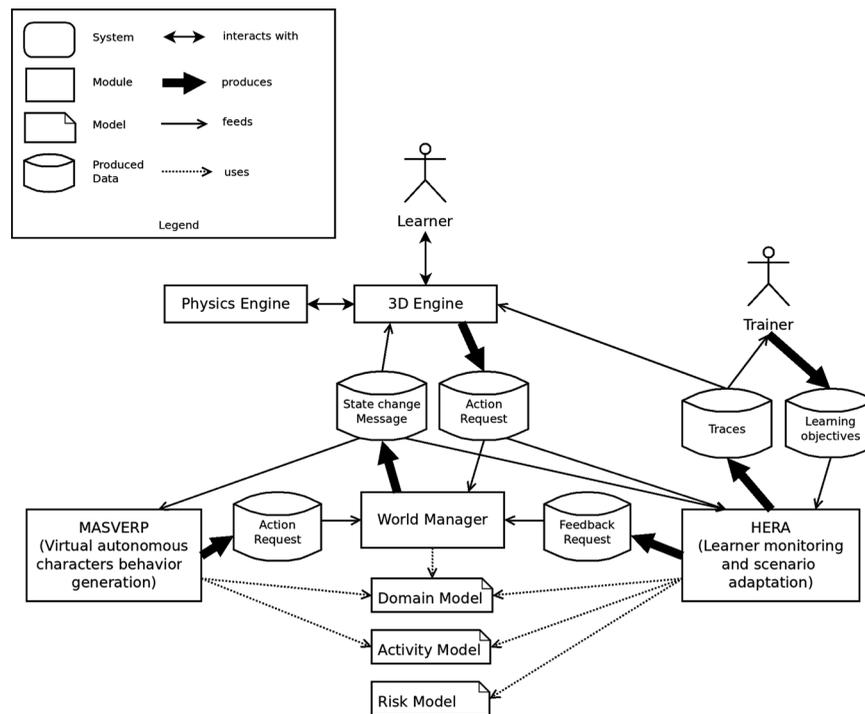


Figure 1. Architecture of the V3S application.

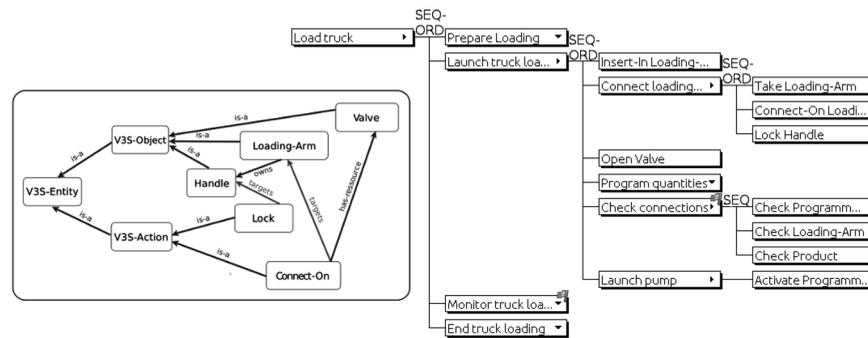
Therefore, instead of opting for a centralized model, the V3S approach was to use common core concepts to link different models together. Using the objects and actions of the domain model, each expert can enter his or her own model explicitly in a graphical formalism, thus providing different views over the same reality. However, even if these models are entered in a graphical editor, they rely on a formal representation that is directly interpretable by computer systems. This approach has two main advantages. First, it enables a partial reuse of the models; for example, the same domain model could correspond to several activity models depending on the subcontracting company that is considered. Secondly, these models are expressed in formalisms that are not computer languages, and thus can be directly entered by experts; this facilitates the integration of human factors, unlike the models from computer scientists that often consider humans as systems.

The process of building a virtual environment involves an iterative and parallel design of these models. The accident analysis helps to identify the technical elements of objects that are considered during decision-making pro-

cesses, thus allowing for the completion of the domain model and the addition of relevant details on the 3D models. The same is true for the activity model, which evolves in an iterative way with the technical system modeling.

**3.1.1 Domain Models.** The domain model is the core model of the simulation. It contains all the objects and associated actions that are required to describe the reference world for the considered domain. These objects, actions, and the semantic relationships between them are described as an ontology. The functioning rules of the technical system are also included in the domain model. They are represented as daemons associated with the concepts of the ontology. An extract of the domain model of an application concerning truck loading is presented in Figure 2.

**3.1.2 Activity Models.** Activity models, or task models, serve as a base for both character behavior generation and learner activity monitoring. Two approaches for these models can be found. On the one hand, human



**Figure 2.** Extract of a domain model and a corresponding activity model.

activity can be expressed through computer formalisms (state machines, Petri networks, neural networks, etc.). These models have to be entered by computer scientists, who are usually not trained in human factors, psychology, or activity analyses. Moreover, in spite of their complexity, these formalisms might not support the expression of the whole range of encountered human behaviors as well as factors that could affect them. On the other hand, psychologists, task analysts, and ergonomists have their own methods for collecting data with high ecological validity, and their own formalisms to describe human activities at work and/or for training purposes (Annett & Duncan, 1967; Annett, Duncan, Stammers, & Gray, 1971). They use, for example, task description languages. The downside is that these languages, being for the most part textual or graphical, are not directly computable and have to be translated into computer formalisms in order to serve as inputs for any module. In addition, even if academic task description languages are numerous, only a few of them are actually used by practitioners (Couix & Burkhardt, 2011). An attempt to create a tool which domain experts, human factors experts, and VR experts can use to build a common activity model was described in Leva et al. (2009). Our approach goes further by linking the activity model to other models of the work situation.

To assist the ergonomists in formalizing these models, we developed a specific activity description language called ACTIVITY-DL (formerly HAWAI-DL; Amokrane, Lourdeaux, & Burkhardt, 2008; Edward et al., 2008). It was inspired by features of two task-description languages developed in the ergonomics

and HCI communities, namely MAD\* (Sebillotte & Scapin, 1994) and GTA (Van Der Veer, Lenting, & Bergevoet, 1996). ACTIVITY-DL aims to support the description of how the operator cognitively represents the task; that is, it is a cognitive rather than logical analysis of the task, in contrast to the HTA (Annett & Duncan, 1967) approach. Such a cognitive fit of the description is related to both the concepts provided by the language features and the methods to collect data about operators' representations of their own activity. The main concepts are a set of possible goals/tasks and subgoals/tasks, which translate the operator's viewpoint of the activity, the relations between these goals, and the possible flow of actions and conditions of their achievement.

The tasks directly reference actions and objects defined in the domain model. The temporal relations between subtasks are specified by constructors which, although they partially keep the same terminology as MAD\* and GTA, are formally defined in terms of temporal Allen relations (Allen, 1983; see Table 1). The preconditions, which define the context in which the task can or has to be executed, are divided into two groups depending on the treatment that is made (either by the user or virtual characters). They are described in Table 2. An extract of the activity model of an application concerning truck loading is presented in Figure 2.

ACTIVITY-DL can be directly interpreted by software modules to generate parts of an operating model depicting constraints and possible actions on objects when attempting to achieve the task within the vir-

**Table 1.** *ACTIVITY-DL Constructors*

Constructors	Name	Allen relations
IND	Independent	$A\{<, >, m, mi, o, oi, s, si, d, di, f, fi, =\}B$
SEQ	Sequential	$A\{<, >, m, mi\}B$
SEQ-ORD	Ordered	$A\{<, m\}B$
PAR	Parallel	$A\{o, oi, s, si, d, di, f, fi, =\}B$
PAR-SIM	Simultaneous	$A\{=\}B$
PAR-START	Start	$A\{s, si, =\}B$
PAR-END	End	$A\{f, fi, =\}B$

tual environment. Since objects and actions are also described in the domain model, the complete operating model results from the interpretation of both information sources, including potential inconsistencies that could exist between the two descriptions.

Moreover, *ACTIVITY-DL* handles the formalization of activity in degraded situations as well as practices resulting from cognitive compromises that could generate risks. It allows the system to deal with errors on two fronts: first, the detection of errors and risk-taking situations in learner activity; second, the demonstration of similar behaviors by the virtual autonomous characters. To this extent, *ACTIVITY-DL* incorporates the concepts of boundary activities tolerated during use (BATUs) and boundary conditions tolerated at use (BCTUs; Fadier, Garza, & Diedlot, 2003). BATUs and BCTUs are concepts derived from research in cognitive psychology, which reflect a local—and often informal—compromise between actors of a certain field. For instance, working with chemicals without wearing individual protection equipment (glasses, gloves, etc.) to save time might exist as a BATU on specific sites; that is, a tolerated risk practice. There are two types of BATUs, corresponding to forbidden (alternatively mandatory) activities that are tacitly allowed to be realized (alternatively not allowed to be realized) by experts under certain conditions.

**3.1.3 Risk Models.** Risk analysis methods can be classified into two categories: classical risk-analysis

methods, and human-reliability analysis methods. The first category allows for a detailed quantitative analysis, but fails to support the complexity resulting from human factors. The second category manages to take into account human factors, but does not provide measures for the system performance.

The MELISSA approach (Camus, Lenne, & Plot, 2012) used in the V3S project combines these two methods. It aims to be the missing link between analysis methods and virtual environment design in the fields of decision-making and risk-management training. Its objective is not to build a complete model of the sociotechnical system, but to help in conceiving scenarios and scenario elements (e.g., relevant object properties), in order to build virtual environments that would be considered coherent by experts regarding risks. Three points of view are considered: (1) activity-related accident scenarios, (2) prescribed task representation, and (3) potential influence factors on the activity. MELISSA brings these together to determine the criteria that would be sufficient for establishing good rules; that is, rules that would ensure efficiency regarding production goals, costs, deadline management, technical specifications, and security goals.

Accident scenarios in the MELISSA method are represented as bowtie graphs. Bowtie graphs are a risk analysis method based on barriers (Hollnagel, 1994). These kinds of approaches consist of identifying all conceivable major accident scenarios, then listing the barriers that oppose the accident development. Risk acceptability depends upon the choices that are made concerning both the number and performance of the barriers that have to be in place in order to consider the risks to be controlled. The central point of the bowtie graph is the critical event. The left part is a failure tree, corresponding to the causes of this event. The right part is an event tree, corresponding to the consequences of this event.

To aid in designing coherent virtual environments, risk analysis has to determine not only the causality chains leading to a particular critical event, but also their occurrence probabilities. Risk occurrence measures can be expressed quantitatively; that is, by the probability, frequency, and/or occurrence rate of a critical event for

**Table 2.** *Preconditions*

Category	Precondition	Description
Conditions to realize	Nomological	Describe the state of the world needed for the task to be physically realizable. Conditions directly depend from the action rules defined in the domain model. Example: To open the door, the door has to be closed.
	Regulatory	Describe the state of the world needed for a good realization of the task according to the prescribed procedure. Example: To unplug the pipe, protective gloves have to be worn.
Conditions to examine	Contextual	Describe the state of the world in which the task is relevant. When these conditions are false, then the task can be ignored. Example: Cleaning the pipe is only relevant if the pipe is dirty.
	Favorable	Describe the state of the world in which the task would be preferred over other ones. These conditions help to choose between several tasks when an alternative exists for the realization of a decomposable task. Example: If the bolt is rusty, unjamming it is preferable rather than directly unscrewing it.

a given population. It can also be expressed qualitatively, using terms such as often, sometimes, rarely, and so on (Polet, Vanderhaegen, & Amalberti, 2003). The risks we consider here can involve human errors, which are a probabilistic and uncertain phenomenon. Not every human error leads to risks, particularly the conscious violations we describe as BATUs. Moreover, risk triggering varies according to the severity of committed errors. Therefore, we completed bowtie graphs with Bayesian networks that allow us to interpret, represent, and compute the occurrence probability of potential risks. The nodes of the Bayesian network correspond to the main concepts of the bowtie graph: initiating events, barriers, critical events, and dangerous phenomena, which are represented in our interpretation as environmental conditions, BATU tasks, independent tasks, and risks, respectively. More information on the translation of bowtie graphs into Bayesian networks can be found in Amokrane and Lourdeaux (2009).

### 3.2 Autonomous Virtual Characters

To expand users' consciousness of the risks they are exposed to, the V3S project aims to represent erroneous behaviors, as well as some of the factors leading operators to them. To this extent, the simulation includes virtual characters that evolve in the virtual environment alongside the user.

Up to now, most of the work conducted on behavior generation for virtual characters leans on frameworks inspired by mathematical formalisms, such as Petri networks or state machines. Indeed, in the cognitivist frame of reference, human reasoning tends to be represented as a computer program. Our aim is not to simulate and explain their reasoning, but to account for the complexity, flexibility, and variability of behaviors that are observed on-site. Therefore, we place ourselves in the situated cognition frame of reference, which implies that cognition is influenced by habits, customs, cultural

context, and so on. To bring human factors into the simulation, our characters show ecological behaviors such as errors and deviations from their procedures.

Because of the complexity of their behaviors, especially in coactivity situations, it would be impossible to control those characters from a global point of view, and therefore they are modeled as autonomous agents. We proposed a multi-agent framework called MASVERP (Multi-Agent System for Virtual Environment for Risk Prevention). This framework is based on the BDI model (Bratman, 1987), enriched by taking into account several physical and physiological characteristics (e.g., thirst, tiredness, stress) and personality traits (e.g., caution, expertise) in the characters' decision processes.

The MASVERP framework integrates several models from cognitive psychology research, including the contextual control model (COCOM) proposed by Hollnagel (1994). The COCOM model is a situated cognition model, which supposes that actions are determined by their context. This context is defined as the interpretation (or understanding) of the situation by the operator; this interpretation can be influenced by knowledge about the goal, and the links between actions, resources, and constraints. The COCOM model thus defines different control modes which depend on the time pressure and influence the anticipation and planning abilities of the characters. The characters select their next actions in the activity model in a similar way to HTN planning (Cavazza, Charles, & Mead, 2002). Depending of their control mode, they more or less adopt the compromising behaviors described by BATUs, and nonintentional errors can be included in their plans. For instance, if the task that the character wants to realize has regulatory conditions and if the character is in strategy mode, then he or she will first try to make these conditions true. However, if the character is in tactical or opportunist mode, he or she will only consider these conditions if their realization does not take much time. In scramble mode, the character will not consider any conditions, and will only do the tasks he or she is used to doing. Nonintentional errors are based on the error taxonomy proposed by Hollnagel in the CREAM model (task omissions, insertions, etc.). A more detailed

description of MASVERP can be found in Edward et al. (2008).

### 3.3 Scenario Adaptation

The V3S project aims at providing coherent, relevant, varied, and efficient training scenarios. Rather than consistently framing the learner's activity, it favors indirect guidance through a nonintrusive adaptation of scenarios. Indeed, enhancing user agency allows for exploratory approaches and trial and error learning, while pedagogical control guarantees learning and ensures its progressiveness. To this extent, the framework includes a learner monitoring and scenario adaptation module called HERA (Helpful agent for safety learning in virtual environments). Its approach differs from other approaches based on learning scenarios in that it supports a learning-by-doing paradigm. Instead of systematically interrupting the learners to explain their errors, it allows them to make errors and to observe the consequences in the virtual environment. Moreover, the dynamic scenario adaptation brings up relevant and progressive training situations.

HERA uses a set of rules to determine the appropriate feedback thanks to its pedagogical model (Amokrane & Lourdeaux, 2010). These rules are based on the learner's level, pedagogical goals, pedagogical situations, risk probabilities, error types, as well as environmental conditions initialized by the trainer at the beginning of a training session. For instance, given that the learner is a novice, that the pedagogical goal is to take risks into account, that the learner does not accompany the valve's lever (BATU error), that the valve's spring is not worn out, and that the risk probability is lower than 0.7, then the feedback would be to send an attention message to the learner. On the other hand, given that the learner is an expert, that the pedagogical goal is to take risks into account, that the learner does not accompany the valve's lever, that the valve's spring is worn out, and that the risk probability is higher than 0.2, then the feedback would be to trigger a leak. HERA's feedback varies according to the learner's level in order, for example, to not disturb novices and to penalize experts, thus fos-

tering a personalized, adaptive, and gradual learning process.

Learners can be provided with three kinds of feedback: situated pedagogical feedback, situated scenario adaptation, and performance criteria. Situated pedagogical feedback can consist of: (1) scale modification; for example, enlarging some parts of an object to get a better view, (2) reification; that is, to show the learner some concepts or abstractions in a concrete and intelligible form, (3) restrictions to limit the learner's actions, such as stop messages sent to novice learners when they commit a severe error, or (4) superposition of information, whether it be comments and argued explanations about the consequences of the learner's actions, or warnings and attention messages sent to attract the learner's attention. Performance criteria are calculated based on learner performance regarding his or her risk behavior, errors, productivity, time management, and so on. As for situated scenario adaptation, it consists of dynamic control of the situation's complexity and of the disruptions of the simulation. The system can either trigger or inhibit the realization of the risks due to learner errors, or send disruptive elements in order to know the reaction of the learner in front of unforeseen situations (e.g., fires, storms, etc.).

### 3.4 Learner Monitoring

To adapt scenarios to the learner's behavior, it is necessary to have an accurate understanding of what he or she is doing. Moreover, especially when the addressed situations are critical, highly complex, and/or need operators to react quickly, it is also necessary to enhance user activity comprehension. The V3S approach on this point is twofold: assist the trainer in his or her monitoring task by providing traces and performance criteria, and help the learner interpret his or her activity, in real time and afterward, through cause-consequence analysis.

Therefore, in addition to scenario adaptation, HERA tracks the learner's activity. However, it does not attempt to track incorrect knowledge for the identification and correction of user misconceptions. It is the learner's responsibility to realize that his or her solu-

tions are wrong for the given situation, and that he or she must experiment with other solutions to understand why. Therefore, the goal of HERA is to detect erroneous actions so that it may modulate their consequences and provide explanations for them. To do this, the learner's activity has to be analyzed in detail. In most monitoring systems based on expertise, learner activity is only compared to the prescribed task, and so any deviation is considered to be an error. Rather than solely considering prescribed activity, we want to compare learner activity with a set of on-site-observed possible behaviors. Moreover, in training, activities that support learning are often far from the prescribed tasks that learners would realize once they have acquired all the knowledge and right practices needed.

To do this, HERA integrates a recognition module. This module uses plan-recognition techniques (Cohen, Perrault, & Allen, 1981) to determine what the learner is doing and detect the errors he or she is committing according to a reference model and an error model. The reference model is described using ACTIVITY-DL (cf. Section 3.1.2), which allows our plan recognition module to detail learners' effective activity by classifying their actions into right practices, errors, and violations. Error classification lies on the error model, which is based on the CREAM model (Hollnagel, 1994), but solely on simple phenotypes (repetition, reversal, omission, delay, intrusion, etc.), as only the learner's observable actions are evaluated. In addition, HERA can go to a higher level of detail by distinguishing between different types of errors: (1) task-related errors (precondition, post-condition, and constructor errors), (2) target object and action errors, (3) BATU errors (i.e., when a safety-related task is performed [alternatively not performed] by the learner while it is forbidden [alternatively mandatory] in the procedure), (4) role errors, or (5) viewpoint errors (i.e., when a task is not performed because of the limitations resulting from the learner's subjective viewpoint). When the learner's errors are related to specific risks, HERA uses the risk model (Section 3.1.3) to determine the occurrence probabilities of those risks, hence ensuring that the simulation stays coherent.

HERA computes every task performed by the learner, every error detected in his or her activity, and every



**Figure 3.** Screenshot of the application.

risk that is incurred because of these errors, in order to produce the learner's trace that can be communicated to the trainer in real time or to the learner at the end of a session. This trace is used to calculate the scenario adaptations to perform on the virtual environment.

### 3.5 Implementation

The V3S project has led to the development of two prototypes. The first one addresses the issue of pipe substitution on chemical-processing sites. The user plays the role of a manager and has to conduct a collaborative procedure while dealing with teammates who can be novice, stressed, tired, and so on. The second prototype is based on a case of loading of hazardous matter on oil depots, where the user is confronted with the consequences of his or her own deviations from the procedure. Figure 3 shows a screenshot of this application.

Both prototypes have been integrated in a photo-realistic 3D environment, developed with 3DVIA Virtools. Furthermore, the second prototype was developed in two distinct versions regarding their technological configurations. The first one is a desktop version, where the user navigates using a mouse and a keyboard, and interacts with objects by using contextual hierarchical pie menus. The second one is an immer-



**Figure 4.** Immersive version of the application.

sive version, using scale-1 stereoscopic visualization, motion capture with the ARTtrack system, and the high-precision physics engine XDE (CEA-LIST, 2012). This version is presented in Figure 4.

## 4 User Study

The evaluations reported in this paper involved the two versions of the second prototype, dedicated to the training of drivers on situations and procedures related to the loading of hazardous material at oil depots. The objectives were threefold:

1. To assess the acceptability of the proposed system.
2. To assess the concrete suitability as well as the relevance of its use in real conditions of training with professional trainers and trainees.
3. To measure the degree of usability of the two (desktop vs. immersive) interface versions.

### 4.1 Method

The methodological approach was adapted to the functional differences between the desktop and immersive versions. Indeed, the desktop version provides the environment and functions required to play real training sessions with trainers and trainees, whereas the immersive version offers a simplified version of the

**Table 3.** Interpretation of the SUS Score in Terms of Usability and Acceptability from Bangor et al. (2008; 2009) adapted with added score ranges

SUS score range	Mean SUS score	Usability adjective scale	Level of acceptability
00.00–44.00	25.00	Worst imaginable	Not acceptable
44.00–51.00	39.17	Poor	Not acceptable
51.00–55.00	52.01	Fair	Marginal low
55.00–75.00	72.75	Good	Marginal high (<70.00) Acceptable (≥70.00)
75.00–87.50	85.58	Excellent	Acceptable
87.50–100.00	100.00	Best imaginable	Acceptable superior

environment and is more oriented toward the evaluation of interface and interaction features. Consequently, the desktop version was evaluated in the context of two real training sessions with real users (the same trainer, 10 trainees) who participate in initial training of tank truck drivers. We carried out an empirical evaluation with the immersive version using a simplified version of the learning scenario—connecting the loading arm to the truck—with a sample of 15 subjects. The experimental setup for the immersive version was composed of a 75-in screen with active stereoscopy, and four infrared cameras arranged at approximately 30 in from each screen corner.

The simplified version of the task included the following subtasks:

1. Select the item to move from a starting point.
2. Move the selected item to the insertion zone indicated as a target.
3. Insert the item in the target area as precisely as possible.

The steps preceding and following these subtasks in the actual procedure—and in the pedagogical scenario of the desktop version—were thus not implemented and used in the immersive version (e.g., manipulation of pumps, pump valves, etc.). Each participant had to achieve the sequence in five sets, each made of six trials. The design and complete details of the procedure and method are beyond the scope of the present paper (see Burkhardt, Repain, Andriot, Brisset, & Niogret, 2011).

## 4.2 Material

Whatever the approach, we used the System Usability Scale (SUS) questionnaire (Brooke, 1996) to evaluate the acceptability and perceived usability of the prototype with the users. The SUS is a simple questionnaire composed of 10 Likert-scale questions, which gives a global view of subjective assessments of usability (cf. Appendix A). It yields a single number representing a composite measure of the overall usability of the system. The standard procedure to compute the SUS score is to sum the score across every item. Each item's score ranges from 0 to 4. For items 1, 3, 5, 7, and 9, the score is the Likert-scale response minus 1. For items 2, 4, 6, 8, and 10, the score is 5 minus the Likert-scale response. Next, we multiply the sum of the scores by 2.5 to obtain the overall value of the SUS score within a range of 0 to 100. The previous analysis of 2,324 surveys from 206 usability tests covering a wide range of interface types (Bangor, Kortum, & Miller, 2008; 2009) demonstrated the reliability of SUS to assess the acceptability of a product/system as well as its predictability of impact of changes on usability. Furthermore, the authors showed that SUS scores can be reliably associated with acceptability and usability adjectives to facilitate the score interpretation and further comparisons (cf. Table 3).

We also used two classes of items from the QUIS (Questionnaire for User Interface Satisfaction) usability questionnaires (Harper & Norman, 1993): terminology and learnability. Specifically for the immersive version,

we also evaluated the subjects' experience of immersion during the test with 11 items adapted from the rating scale by Jennet et al. (2008) to assess the experience of immersion in games and virtual environments.

### 4.3 Collected Data

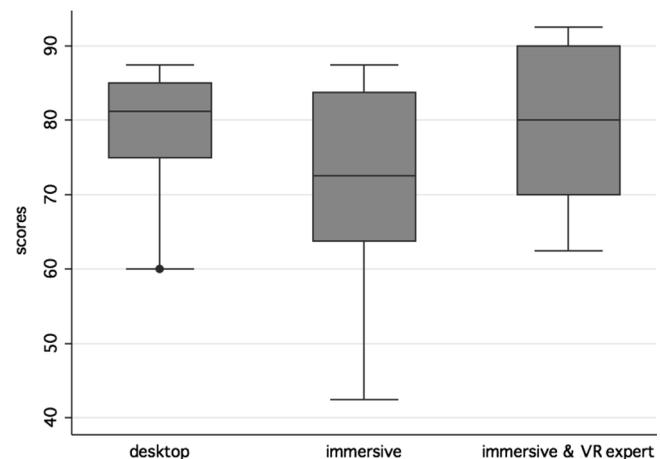
We recorded the sessions and conducted interviews with participants in both the real training sessions and the experiment with the immersive version. Furthermore, performance data on the simplified task were automatically recorded during the experiment.

### 4.4 Results

This section provides the main results of the evaluations. SUS scores in both conditions are presented in Figure 5.

**4.4.1 Field Evaluation.** Regarding the evaluation in the field with real trainees, questionnaire results showed an average SUS score of 81/100 ( $SD = 10$ ) which can be considered as an indication that this version is acceptable with an excellent usability rating. The mean score for the QUIS items concerning the terminology was 7.91/9 ( $SD = 1.22$ ) and the mean score for learnability was 7.71/9 ( $SD = 1.3$ ). These scores are far higher than 5/9 which has been proposed as arbitrary value indicating a mediocre level of usability (cf. notes in QUIS site at <http://lap.umd.edu/quis/QuantQUIS.htm>). An item-by-item analysis indicated two dimensions with a slightly lower score: use of terms and quality messages, suggesting possible improvements in the interface of the desktop version.

More qualitatively, interviews with trainers and managers in the training company showed that they were highly satisfied with the way that the demonstrator could fit in their initial training sessions. They proposed to extend the use of the virtual environment to more specialized professional courses, including emergency intervention training. In addition, most of the trainees spontaneously mentioned their interest in accessing an online individual version of the vir-



**Figure 5.** SUS scores in both conditions and depending on whether the subjects were familiar with VR.

tual environment in order to go on with their training at home.

In addition, a third informal observation was performed using the desktop version as a basis for a role-play-like training session, where three trainers were in charge of putting one individual learner on the spot and disrupting him during the simulation. The truck driver made numerous mistakes, which he explained by the fact that he was panicked and rushed. Afterward, the trainers praised the use of the device to train drivers to act in stressful situations.

**4.4.2 Immersive Version Evaluation.** Performance data in the immersive version showed that subjects achieved the task in about 30 s ( $SD = 25$  s) with high variability. A significant learning effect of trials was also observed,  $F(1, 14) = 5.33$ ,  $p < .0368$ , as subjects required less time to achieve the task when the number of trials increased.

Regarding the usability and acceptability of this immersive version, the results show it is also acceptable but at a slightly lower score (average SUS score of 75/100,  $SD = 13$ ). For example, participants reported that the interfaces were sometimes perceived as invasive, particularly for the infrared markers. Subjects with video game and virtual reality system experience tended to rate acceptability at a higher level (SUS score = 79/100) than those without such previ-

ous experience (SUS score = 71/100). However, the difference failed to be significant,  $t(13) = \sim -1.1158$ , n.s., which could be due to the low number of subjects. It should also be noted that the score is descriptively lower than in the desktop version. This result might also be explained by the lower level of function of the prototype and the simplified nature of the experimental situation, as well as by the fact that participants were not real trainees and trainers. The specific measure of subjective experience of immersion showed to be moderate (immersion subjective rating = 67/100,  $SD = 5.60$ ). Subjects with a previous experience of virtual games exhibited a slightly lower score ( $m = 66$ ,  $SD = 4.97$ ) than the newcomers ( $m = 68$ ,  $SD = 6.30$ ). However, the effect failed to be significant,  $t(13) = 0.6808$ , n.s.

## 5 Discussion

The results from the evaluations conducted on the prototype raise the possibility of using it in genuine training sessions at a larger scale. For now, only some of the features have been tested with end users, but the prospects of using the demonstrator as a support in other learning situations already reveal new needs.

The evaluations conducted with the prototype only focused on the aspects of utility, usability, and acceptability. It would also be relevant to assess the impact of the different configurations on the learning process, especially for the immersive version that embodies natural interaction in an entirely physical environment. Similarly, it would be necessary to study the impact on the user of the virtual autonomous characters and to estimate the contribution of the progressive learning scenario. Moreover, since the evaluation of the immersive version of the prototype showed that subjects found markers to be a potential shortcoming both for acceptability and for the subjective experience of immersion, the development of a new version has started, using Microsoft's Kinect to perform motion capture without the need for markers or other invasive devices. It would be interesting to evaluate this version, as the invasiveness of the infrared markers was criticized in the previous version.

The V3S framework seems to fulfill the objective of being generic, as it has been used for the creation of three prototypes, differing in either their application cases, technological settings, or pedagogical strategies. Indeed, as for technological settings, it has been used with different input devices (mouse, ART tracking, Kinect camera) as well as different displays (computer screens, 3D TV, wall). Concerning pedagogical strategies, our prototypes on the loading of hazardous matter have been used in real training sessions both:

- As a course material: the trainer demonstrating to the class a virtual site visit, and making learners interact one-by-one for demonstration purposes.
- For tutoring sessions: a role-playing session with one trainee and three trainers focused on putting the trainee in stressful conditions.

Even if no evaluation were conducted in such settings, we also think that this system could be used for supervised group activity, by extending the learner monitoring module to follow several learners simultaneously, as well as for self-training, with pedagogical feedback being sent to the learner according to his or her actions. In addition, the pedagogical adaptability of the system is increased by the different forms of feedback that can be provided to the learner, including the adaptation of the consequences of his or her actions.

To address the new pedagogical situations suggested by the training company, we propose to extend the work done on scenarios within the V3S project. Currently, scenario control mostly involves the triggering or inhibition of events related to the simulated technical system (e.g., leaks, tank overflows). However, it does not take into account the behavior of the autonomous virtual characters evolving in the environment. Indeed, the module managing their behavior and the one responsible for the scenario adaptation part are completely independent of each other. The characters progress on their own, with no interaction with the user and no adaptation to his or her activity, whatsoever. We plan to link these modules so that characters could react appropriately to user actions, depending on the pedagogical objectives of the training sessions, pedagogical rules, user level, traces from previous sessions, and so on.

These characters might then act to maximize or minimize the risks associated with a collaborative procedure, to assist the learner when in trouble by providing advice or performing the task, or, on the contrary, to harass or disrupt the learner to put him or her under stress when the task seems too easy.

A more global scenario adaptation would allow us to introduce the more complex training situations for which training companies have expressed a need. In emergency intervention training, for instance, the teacher should be able to exert control over the events in the virtual environment, and to guide the learner's actions somehow to engender relevant learning situations (a leak, a fire, a storm, etc.). In addition, the transposition of role-playing onto virtual world characters would benefit the trainer, avoiding the need to involve several individuals for a given training session. It could also increase the motivation of the trainee in situations of self-training. By directing scenarios, drawing inspiration from the interactive storytelling field in terms of narratives and enriching them with pedagogical control, we could create richer environments where the unfolding of events would adapt to the learners' level and activity.

The main challenge we face for the introduction of these adaptive scenarios is related to the fact that our environments contain autonomous entities, whether they be human users or virtual characters. Moreover, since we foster a learning-by-doing paradigm, it is vital to maintain a high level of user agency. Trade-offs must therefore be established, on one hand between the user's feeling of freedom and the respect of the desired course of events, and on the other hand between the virtual characters' autonomy and the preservation of scenario consistency.

## 6 Conclusion

To address the risk-management training issue, the V3S consortium proposed a generic framework for virtual environments that was used to create VEs adapted to different pedagogical and technological settings, and used them in genuine training contexts. This framework

uses generic description languages and decision engines that interpret models based on field analysis and taking into account human factors. The use of ergonomic and risk analyses helps to increase the ecological validity of the environment, through the generation of non-ideal virtual characters' behaviors whose performance is affected by external and internal factors such as stress, and characters who are prone to compromise depending on their level of expertise. The pedagogical control of the simulation is provided by a learner-monitoring module, which is able to adapt the reactions of the environment to ensure a gradual training process.

The results from the V3S project already meet many needs of initial or continuing training in risk management. The V3S project led to the development of the HUMANS platform (Lhommet, Lourdeaux, & Barthes, 2012), a generic software platform for VEs based on models of humans at work. The multiplicity of uses of virtual environments for professional training now offers new perspectives, whether in the use of immersive devices to interact with the simulations or in the management of adaptive scenarios underlying them. In this event, the HUMANS platform will integrate an orchestration module called SELDON, dedicated to dynamic and coherent adaptation of the simulation scenario.

## Acknowledgments

The authors want to thank the ANR/RNTL for the funding of the V3S project, as well as all the members of the consortium: UTC/Heudiasyc UMR7253, INERIS, Paris Descartes University, CEA-LIST, Emissive, EBTRANS, CICR, SI-GROUP, TICN, and APTH. This work has been partially funded by the Picardy region and the European Regional Development Fund (ERDF) under the FIANNA and ARAKIS projects.

## References

- Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11), 832–843.

- Amokrane, K., & Lourdeux, D. (2009). Virtual reality contribution to training and risk prevention. *Proceedings of the International Conference of Artificial Intelligence, ICAI '09*, 726–732.
- Amokrane, K., & Lourdeux, D. (2010). Pedagogical system in virtual environments for high-risk sites. *Proceedings of the Second International Conference on Agent and Artificial Intelligence*, 371–376.
- Amokrane, K., Lourdeux, D., & Burkhardt, J. (2008). HERA: Learner tracking in a virtual environment. *International Journal of Virtual Reality*, 7(3), 23–30.
- Annett, J., & Duncan, K. D. (1967). Task analysis and training design. *Journal of Occupational Psychology*, 41, 211–221.
- Annett, J., Duncan, K. D., Stammers, R. B., & Gray, M. J. (1971). *Task analysis*. London: Her Majesty's Stationery Office.
- Bainbridge, L. (1987). Ironies of automation. In J. Rasmussen, K. D. Duncan, & J. Leplat (Eds.), *New technology and human error* (pp. 271–284). Chichester, UK: Wiley.
- Bangor, A., Kortum, P., & Miller, J. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction*, 24, 574–594.
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, 4, 114–123.
- Bell, J. T., & Fogler, H. S. (2000). *A virtual reality safety and hazard analysis simulation*. Paper presented at the conference of the American Society for Engineering Education, St. Louis, MO.
- Bell, J. T., & Fogler, H. S. (2004). The application of virtual reality to (chemical engineering) education. *Proceedings of IEEE VR 2004*, 217–218.
- Besnard, D., & Cacitti, L. (2005). Interface changes causing accidents. An empirical study of negative transfer. *International Journal of Human–Computer Studies*, 62, 105–125.
- BinSubaih, A., Maddock, S., & Romano, D. (2006). A serious game for traffic accident investigators. *International Journal of Interactive Technology & Smart Education*, 3, 329–346.
- Bratman, M. E. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). London: Taylor & Francis.
- Burkhardt, J. M., Bardy, B., & Lourdeux, D. (2003). Immersion, réalisme et présence dans la conception et l'évaluation des environnements virtuels. *Psychologie Française*, 48, 35–42.
- Burkhardt, J.-M., Lourdeux, D., & Lequatre, F. (2005). Environnements virtuels pour l'apprentissage: De l'image d'épinal à la réalité des usages et des configurations socio-techniques. *ACM IHM 2005*, 163–170.
- Burkhardt, J.-M., Repain, D., Andriot, C., Bricquet, J., & Niogret, J.-L. (2011). *Evaluation de l'acceptabilité et de l'utilisabilité des démonstrateurs V3S en version immersive et non immersive* (Report No. V3S ANR project). Paris: Université Paris Descartes.
- Camus, F., Lenne, D., & Plot, E. (2012). Designing virtual environments for risk prevention: The MELISSA approach. *International Journal on Interactive Design and Manufacturing*, 6(1), 55–63.
- Cavazza, M., Charles, F., & Mead, S. J. (2002). Planning characters' behaviour in interactive storytelling. *The Journal of Visualization and Computer Animation*, 13, 121–131.
- CEA-LIST. (2012). *Xde*. Retrieved from <http://www.kalisteo.fr/lsi/en/aucune/a-propos-de-xde>.
- Cohen, P., Perrault, C., & Allen, J. (1981). Beyond question-answering. *DTIC Research Report ADA100432*. Fort Belvoir, VA: Defense Technical Information Center.
- Couix, S., & Burkhardt, J.-M. (2011). Task descriptions using academic oriented modelling languages: A survey of actual practices across the SIGCHI community. *Proceedings of the 13th IFIP TC13 Conference on Human–Computer Interaction*, 555–570.
- Delmas, G., Champagnat, R., & Augeraud, M. (2007). Plot monitoring for interactive narrative games. *Proceedings of the International Conference on Advances in Computer Entertainment Technology, ACE '07*, p. 17.
- Edward, L., Lourdeux, D., Lenne, D., Barthes, J. P., & Burkhardt, J. (2008). Modelling autonomous virtual agent behaviours in a virtual environment for risk. *IJVR: International Journal of Virtual Reality*, 7(3), 13–22.
- Fabre, D., Couix, S., Burkhardt, J.-M., Gounelle, C., & Cabon, P. (2006). Virtual reality to support human factors for safety: Where we are and where we (aim to) go. *ESREL, Safety and Reliability Annual Conference*, Estoril, Portugal.
- Fadier, E., Garza, C. D. L., & Didelot, A. (2003). Safe design and human activity: Construction of a theoretical framework from an analysis of a printing sector. *Safety Science*, 41(9), 759–789.

- Gardeux, F., Marsot, J., & Rolin, A. (2005). Un environnement virtuel pour la formation à la prévention des risques professionnels. *1st International VR-Learning Seminar*, Laval, France.
- Gounelle, C., Cabon, P., Burkhardt, J. M., Couix, S., Fabre, D., Anastassova, M., et al. (2007). *Integrating human factors approaches with virtual reality for safety in the VIRTUALIS project*. Paper presented at the Virtual Reality International Conference, VRIC, Laval, France.
- Haller, M., Kurka, G., Volkert, J., & Wagner, R. (1999). OMVR, A safety training system for a virtual refinery. *Proceedings of ISMCR'99, Topical Workshop on Virtual Reality and Advanced Human-Robot Systems*, 291–298.
- Harper, B., & Norman, K. L. (1993). Improving user satisfaction: The Questionnaire for User Interaction Satisfaction Version 5.5. *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference*, 224–228.
- Hollnagel, E. (1994). *Human reliability analysis: Context and control* (1st ed.). San Diego, CA: Academic Press.
- Jamieson, G. A., & Vicente, K. J. (2001). Ecological interface design for petrochemical applications: Supporting operator adaptation, continuous learning, and distributed, collaborative work. *Computers and Chemical Engineering*, 25, 1055–1074.
- Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., & Tijs, T. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66, 641–661.
- Johnson, C. (1998). Evaluating the contribution of desktop VR for safety critical applications. *18th International Conference on Computer Safety, Reliability and Security, SAFECOMP '99*, 67–78.
- Leva, M., Kay, A., Mattei, F., Kontogiannis, T., De Ambroggi, M., & Cromie, S. (2009). A dynamic task representation method for a virtual reality application. *Engineering Psychology and Cognitive Ergonomics*, 32–42.
- Lhomme, L., Lourdeaux, D., & Barthes, J.-P. (2012). Foule sentimentale: Influence des caractéristiques individuelles sur la contagion émotionnelle dans une foule informatique. *Revue d'Intelligence Artificielle*, 26(3), 281–308.
- Li, L., Zhang, M., Xu, F., & Liu, S. (2006). ERT-VR: An immersive virtual reality system for emergency rescue training. *Virtual Reality*, 8, 194–197.
- Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, 31, 557–564.
- Magerko, B., Wray, R., Holt, L., & Stensrud, B. (2005). Improving interactive training through individualized content and increased engagement. *The Interservice/Industry Training, Simulation & Education Conference, I/ITSEC*.
- Mollet, N., & Arnaldi, B. (2006). Storytelling in virtual reality for training. In Z. Pan, R. Aylett, H. Diener, X. Jin, S. Glöbel, & L. Li (Eds.), *Technologies for e-learning and digital entertainment* (Vol. 3942, pp. 334–347). Berlin: Springer-Verlag.
- Polet, P., Vanderhaegen, F., & Amalberti, R. (2003). Modelling borderline tolerated conditions of use (BTCU) and associated risks. *Safety Science*, 41(2–3), 111–136.
- Sebillotte, S., & Scapin, D. L. (1994). From users' task knowledge to high-level interface specification. *International Journal of Human-Computer Interaction*, 6, 1–13.
- Shawver, D. (1997). Virtual actors and avatars in a flexible user-determined-scenario environment. In *Proceedings of the IEEE Virtual Reality Annual International Symposium*, 170–171.
- Stytz, M. R., Banks, S. B., Garcia, B. W., & Godsell-Stytz, G. M. (1997). A virtual environment for emergency medical training. *Proceedings of the 6th International Conference for Man-Machine Interaction Interfaces 97*, 214–216.
- Swartout, W., Gratch, J., Hill, R. W., Hovy, E., Marsella, S., Rickel, J., et al. (2006). Toward virtual humans. *AI Magazine*, 27, 96–108.
- Ustarroz, A., Lozano, A., Matey, L., Siemon, J., Klockmann, D., & Berasategi, M. I. (2004). VIRTUOL, Virtual Reality for Machine-Tool Training. *Mécanique & Industries*, 5, 207–212.
- Van Der Veer, G., Lenting, B., & Bergevoet, B. (1996). GTA: Groupware task analysis—Modeling complexity. *Acta Psychologica*, 91, 297–322.

**Appendix A: Items of the SUS Questionnaire (Brooke, 1996)**

---

		Strongly agree	Strongly disagree
1.	I think that I would like to use this system frequently		
2.	I found the system unnecessarily complex		
3.	I thought the system was easy to use		
4.	I think that I would need the support of a technical person to be able to use this system		
5.	I found the various functions in this system were well integrated		
6.	I thought there was too much inconsistency in this system		
7.	I would imagine that most people would learn to use this system very quickly		
8.	I found the system very cumbersome to use		
9.	I felt very confident using the system		
10.	I needed to learn a lot of things before I could get going with this system		

---