Scenario-based Training: Director's Cut

Marieke Peeters[1,2,4], Karel van den Bosch[2], John-Jules Ch. Meyer[1,2], and Mark A. Neerincx[2,3]

[1] Information and Computing Sciences, Utrecht University
[2] Training Innovations, TNO - Human Factors
[3] Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
[4] Corresponding author: mpeeters@cs.uu.nl

Abstract. Research regarding autonomous learning shows that freeplay does not result in optimal learning. Combining scenario-based training with intelligent agent technology offers the possibility to create autonomous training enriched with automated adaptive support delivered by a director agent. We conducted an experiment to investigate whether *directing* training scenarios improves the quality of training. Six instructors rated video fragments of directed and non-directed scenarios in terms of learning value. Results show that the instructors consider directed scenarios to be considerably more effective for learning than non-directed scenarios. Implications for the design of a director agent are discussed.

Keywords: intelligent agents, autonomous training, director agent

1 Introduction

Scenario-based training (SBT) is a powerful way to let trainees prepare, execute, and evaluate real (authentic) tasks within a simulated environment [4, 12]. SBT meets the principles recognized in dominant instructional theories as described by Merrill (2002) [9]. Important benefits of training within a simulated environment are the reduction of risks and the possibilities for control over training, e.g. authoring the scenario, delivering feedback, and instructing the actors. However, this control can only be exerted when the scenario is not playing. Control *while* the scenario unfolds is problematic, if not impossible. Yet such control is also highly desirable. Research has shown that trainees need a suitable amount of support *during* training tasks [5]. For instance, if the trainee is performing well, it would be interesting to tell an actor to make a mestake. Whereas if the trainee panicks, it would be better to tell an actor to take over. During normal SBT, such adjustments are hard to accomplish. However, by using intelligent agent technology it becomes possible to wield *online control* over training in advanced practice environments, such as serious games [3]. This can be achieved by developing a *director agent* (DA) that controls the scenario as it unfolds; it monitors the course of events in the training environment, analyzes and assesses suitable ways to proceed, and instructs non-player characters (NPCs) to execute, or refrain from, particular actions. The DA uses its means of control to create meaningful and suitable experiences for the trainee.

1.1 Automated Control: the Director Agent

The issue of this paper is how to automate control over a training scenario as a means to guide and support the trainee. The idea to obtain control over a scenario while it unravels, is not new [2, 16]. Within the domain of interactive narrative, there are interesting publications on this subject. In several papers the concept of a director agent (DA) is mentioned, and whereas some researchers merely describe an architecture [7], others actually implemented a framework or built a prototype [8, 10, 14]. Within the mentioned paradigm, the reason for an intervention is a narrative discrepancy, e.g. the player (Little Red Riding Hood) decides to visit her grandmother by bicycle, therefore, the DA intervenes to hold on to the original storyline by giving the player a flat tire.

The current paper will focus on a different reason for intervening, i.e. to create learning opportunities for the trainee that lie within the zone of proximal development [11, 15]. Such opportunities are challenging, yet not confusing [13], but most certainly not boring [1]. This paper focuses on such *pedagogical* interventions. During SBT, instructors use their experience and intuition to intervene; they recognize that a trainee seems lost, overwhelmed or bored and decide to adjust the scenario to attune it to the trainee's needs. To be able to automate these interventions, we need to turn such implicit notions into explicit ones, for instance by defining behavioral cues and events that accompany confusion or boredom, e.g. a lack of activity, the amount of mestakes, posture, etc.

Pedagogical interventions can be divided into two types: supportive and challenging interventions. Supportive interventions are needed when the trainee is performing actions leading him to a situation that is too complex. The trainee receives support to get through some overly complex situation, while leading him to a less demanding situation. Challenging interventions are executed when the trainee is performing all the right actions, but is not being sufficiently challenged. The trainee is motivated to take the training to a higher level. Interventions can consist of adjustments of the complexity level, the availability of information, the salience of certain cues or the amount of learning goals addressed simultaneously.

But even if we define such explicit cues for interventions, the question still remains how effective such interventions are. Clearly, the goal of the interventions is to improve the quality of learning. We argue that a learning situation offers optimal learning opportunities if a trainee is able to cope with the demands, while still being challenged to learn new things [11]. The proximity of a training situation to this optimum can be expressed as the *learning value*. If a training situation has a low learning value, this means that the situation does not meet the trainee's needs: the trainee is either incapable of coping with the demands or he is not being challenged enough to motivate him. In both cases an intervention would be necessary to attune the scenario to the trainee's needs. The question is: Do interventions actually lead to an improvement of the learning value?

Research Question and Hypotheses. The research question in the current study is: "*Will the director's interventions during scenario-based training, triggered by explicit behavior cues, improve the learning value of the scenario?*"

We hypothesize that interventions of a director will improve the learning value of the training scenario as rated by professional instructors.

Chosen Task Domain: 'Bedrijfshulpverlening'. We chose 'bedrijfshulpverlening' (BHV, a Dutch word) to be the task domain for our research. BHV entails the application of first aid and fire fighting by a team of company employees. We created four scenarios: (A) a diabetic woman suffering from hypoglycemia, (B) a lady trapped within a room because of a fire in a trash can near the door, (C) an unconscious cleaning lady, who fainted because of an intoxicating gas and (D) a woman with a broken hip (as a result of fleeing in panic from a fire) lying near a fire hazard. Scenarios were developed to train one individual BHV member. All scenarios included two NPCs playing the roles of victim and bystander.

A detailed script enabled the director to intervene in the scenario in predefined ways. Supportive as well as challenging interventions were triggered by possible behaviors of the trainee. For example, the director's script for scenario (A) contained the following line: "If the trainee is asking irrelevant questions for over three minutes (behavioral cue), the victim is instructed to tell the trainee that her vision is blurred (intervention)." Other cues for supportive interventions included: the trainee repeatedly calls emergency services or fails to perform certain checks. The director used these cues to initiate supportive interventions, e.g. instructing the NPCs to reassure the trainee or to offer their assistance.

A challenging intervention was triggered if the situation proved to be too simple for the trainee to handle, indicated by perfect or near perfect performance. The following rule comes from the director's script of scenario (B): "If the trainee communicates his plans and checks the door of the burning office according to protocol (cues), the bystander is instructed to remain passive (intervention)." Examples of behavioral cues for challenging interventions included: making eye contact with bystander and victim, remaining calm, and giving clear instructions. The resulting challenging interventions included instructing the NPCs to: ask for trainee's attention simultaneously, withhold important information, or create extra complications (e.g. running into a fire hazard).

Prototype: Wizard of Oz Set-Up. Because of the laborious task of implementing a prototype of the envisioned training system, we developed a Wizard of Oz prototype; all agents (NPCs and director) were human and the simulated environment was not virtual. All scenarios took place within a real office room at trainees' company building. This gave us the opportunity to investigate approaches for directing training and their effects on the quality of training.

Two NPCs (human actors), playing the roles of bystander and victim, both received two versions of the behavior they were to display during the scenarios: a supportive and a challenging version. Supportive behaviors were helpful to the trainee. Challenging behaviors were impeding or distracting. Another script was developed for the director. This script contained explicit trainee behavior cues, triggering the director to intervene in specific ways while the scenario unfolded. The execution of an intervention was implemented by instructing the actors to change their behavior from supportive to challenging or vice versa.

2 Methods

2.1 Raters

Six experienced instructors in BHV were asked to rate the video-fragments.

2.2 Materials

Footage. We selected twenty video fragments as a test set. Each fragment contained a part of a recording of a trainee playing one of the aforementioned BHV scenarios. All selected video fragments contained trainee behavior cueing an intervention. In half of the fragments shown to the instructors, the director executed all the interventions (*directed condition*) by telling the actors through in-ear portophones to switch between their behavior variations. In the other half of the fragments, the director was absent; even though the fragments all contained behavioral cues, the associated interventions were not executed (*non-directed condition*). Additionally, both conditions (directed and non-directed) contained five fragments that started off with the actors playing their supportive parts (*supportive startup*), and five fragments that started off with the actors playing their challenging parts (*challenging startup*).

Questionnaire. The raters were asked to evaluate the learning value of the situation for a particular trainee by answering the following question.

”*The learning situation at this point in time offers the trainee . . . opportunities to achieve the learning goals at his own level.* ”

-3	-2	-1	0	1	2	3
absolutely no	no	not really any	maybe some	some	enough	exactly the right

2.3 Procedure

The raters received an elaborate instruction to this experiment, containing an explanation of scenario-based training, exemplified by a video fragment. The four scenarios were explained and the learning goals of each scenario were explicitly pointed out. Finally the raters received instructions regarding the procedure of the experiment and explanations to the questionnaire. The raters were oblivious of the research question of the experiment.

Raters were then presented with two sets of video fragments (a practice set and a test set) following a standard procedure. The video fragment was introduced by a short description of the original scenario and the intended learning goal. The part of the fragment preceding the point of intervention was shown. At the cue for intervention, the fragment was paused and the raters were asked to rate the learning value (rating moment 1). Subsequently, the fragment was continued and paused again at the time the result of the intervention (or the

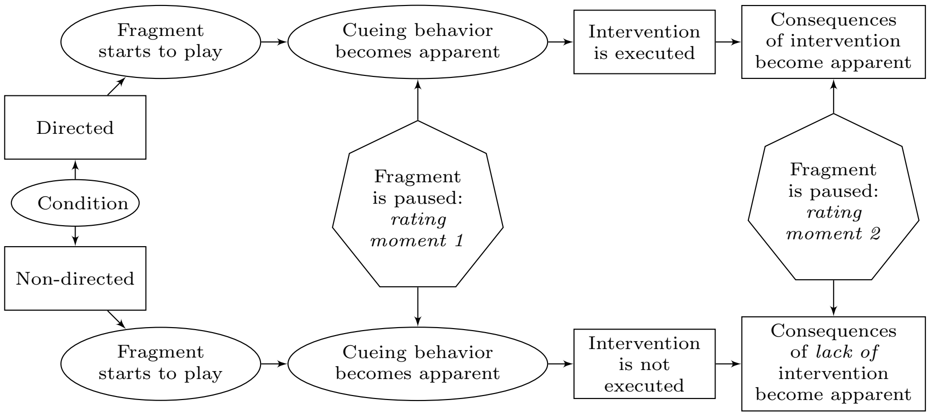

Fig. 1. A graph of the procedure during the experiment

lack thereof) became apparent. The raters were again asked to rate the learning value (rating moment 2). A diagram of the procedure can be found in Fig. 1.

To test and enhance agreement among raters, they were presented with a practice set of 16 video fragments. The raters were encouraged to discuss their judgments in between series to reach consensus on how to value a learning situation. After the practice set, the experiment proper started, by presenting the test set consisting of twenty video fragments to the raters. The raters were not allowed to discuss their judgments, nor could they see each other's judgments. After the test set, the raters participated in a group discussion about their experiences with scenario-based training and their opinions about the video fragments.

2.4 Analysis

An intra-class correlation analysis was performed to assess inter-rater reliability. A repeated measures ANOVA was used to compute the effects of direction upon the rated learning value of the scenario.

3 Results

Data Exploration and Inter-rater Reliability. Forty ratings per rater (two rating moments for a total of twenty fragments) were entered into the analysis. The consistency intra-class correlation coefficient was 0.694 for average measures ($p<.001$). An inter-rater agreement between 0.60 and 0.79 is considered substantial [6], therefore we consider these data to be appropriate for further analysis.

Repeated Measures Analysis. In order to test whether the interventions of the director had an effect on learning value, rated learning values were entered

Table 1. results of the repeated measures analysis *) p <.05 **) p <.01 [1]) one-tailed

effect	F	effect size (partial η^2)	power
director (presence vs absence)[1]	13.847**	.735	.841
startup variation (supportive vs challenging)	11.043*	.688	.757
director (presence vs absence) * rating moment[1]	27.339**	.845	.984

into a repeated measures analysis with two independent factors: director (presence vs absence) and startup variation (a scenario starting in the supportive vs challenging behavior variation). The results of this analysis are shown in Table 1.

A main effect of direction was found (F(1,5)=13.85; p<.01, one-tailed). Examination of this effect showed that the directed fragments received a significantly higher learning value (M=1.08; SE=.31) than the non-directed fragments (M=.35; SE=.23). A second main effect showed a significant difference between the learning value assigned to the two startup conditions (F(1,5)=11.04; p<.01, two-tailed). Overall, the video fragments in the supportive startup condition received a higher learning value (M=.98; SE=.31) than those in the challenging startup condition (M=.45; SE=.22).

Our main interest is the effect of an intervention on the situation's learning value. Therefore the differences between the director conditions (present vs absent) at rating moment 2 are of importance. It is expected there are no differences between the two conditions at rating moment 1. A significant interaction effect between director (presence vs absence) and rating moment (prior to vs after the cue for intervention) (F(1,5)=27.34; p<.01, one-tailed test), showed that indeed there was no significant difference between the directed and the non-directed condition at rating moment 1 (M=.60 vs M=.43, respectively). However, if an intervention was executed at rating moment 2 (director present), the learning value was significantly higher than when no intervention had taken place (director absent) (M=1.55 vs M=.27, respectively). The means belonging to this interaction effect can be found in the row 'overall' of Table 2.

Table 2. mean rated learning value (SE) *)p <.05, one-tailed

	director present		director absent	
	moment 1	moment 2	moment 1	moment 2
challenging startup	.433 (.336)	1.467* (.470)	.233 (.285)	-.333 (.276)
supportive startup	.767 (.391)	1.633* (.363)	.633 (.336)	.867 (.418)
overall	.600 (.306)	1.550* (.394)	.433 (.262)	.267 (.324)

To find out whether the beneficial effect of the director's interventions is equal for both directions of interventions (from supportive to challenging or vice versa), one-tailed 95% confidence intervals of the means were computed for both startup conditions. The interaction effects were significant (p<.05, one-tailed) for both directions of intervention, (see also Table 2), although the effect was stronger for supportive interventions (changing the actor behavior from challenging to supportive).

4 Discussion

The goal of the present study was to investigate the effects of interventions upon the learning quality of a scenario. We created scripts for a director specifying when and how to intervene. Interventions consisted of adaptations in the behavior of the actors (NPCs) and were implemented on-line, i.e. while the scenario unfolded. Video recordings of directed and non-directed training scenarios were shown to experienced instructors, who were asked to rate the learning value of the presented situations. Instructors were naive with respect to the purpose and design of the experiment.

Results confirmed our hypothesis. The rated learning value of scenarios that proceed undirected, without adaptation, were at a fairly low level both halfway and at the end of the scenario. In contrast, the learning quality of directed scenarios improved significantly as a result of the interventions directing the actors to behave appropriately to the performance level of the trainee. Thus, overall, interventions improve the learning value of scenarios. If we examine these results more closely, split for supportive and challenging startup conditions, it becomes clear that scenarios that started in the supportive mode also offer some learning opportunities in the absence of a director. Even though the trainee could use an extra challenge, the mere practice of already acquired skills is still considered useful. However, in the directed condition, it becomes possible to create an extra challenge for the trainee, which results in an even higher learning value. A different pattern is found for the scenarios that started in the challenging mode. For these scenarios, the learning value drops dramatically over time when there is no director present to adjust the scenario. However, in the presence of the director, support is given to the trainee, thereby most likely saving the trainee from losing track and motivation and increasing the learning value of the training.

In a group interview conducted after the experiment, we explained the purpose and design of the study to the instructors and asked them for their experiences in their everyday work. The instructors stated that they find it hard to successfully intervene once they notice that a scenario loses track. They argue that they do realize it when a training situation requires intervention, but that they find it hard to specify beforehand what cues indicate this need. A more practical problem that they put forward is that - in their experience - participating actors tend to be unaware of what is needed, and that it is difficult for instructors to bring across appropriate adjustments to the actors while the scenario is playing. Instructors therefore consider it important to have appropriate and practical instruments to execute the necessary control over their training scenarios. They added to welcome this type of studies to accomplish this need.

In this study we explicitly described cues based upon trainees' responses to specify different types of interventions. These interventions proved to be beneficial to the learning value of the scenario. A next step would be to further refine the different types of interventions a director can execute and to conceptualize the knowledge that is needed to implement such interventions. In the end, the goal is to develop automated systems that formalize relationships between events, learning objectives, trainee behaviors and NPC behaviors to create autonomous, adaptive and effective training scenarios.

Acknowledgements. The authors would like to thank Stichting BHV Nederland for their cooperation in this experiment.

References

1. Baker, R.S.J., D'Mello, S.K., Rodrigo, M., Mercedes, T., Graesser, A.C.: Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three computer-based learning environments. Int. J. Hum-Comp. St. 68(4), 223–241 (2010)
2. Blumberg, B., Galyean, T.: Multi-level Control for Animated Autonomous Agents: Do the Right Thing... Oh, Not That.. In: Creating Personalities for Synthetic Actors, Towards Autonomous Personality Agents. pp. 74–82. Springer-Verlag (1997)
3. van den Bosch, K., Harbers, M., Heuvelink, A., van Doesburg, W.: Intelligent agents for training on-board fire fighting. In: 2nd International Conference on DHM. pp. 463–472. Springer (2009)
4. Cannon-Bowers, J., Burns, J., Salas, E., Pruitt, J.: Making Decisions Under Stress: Implications for Individual and Team Training, chap. Advanced technology in scenario-based training, pp. 365–374. Washington DC: APA (1998)
5. Kirschner, P., Sweller, J., Clark, R.: Why minimal guidance during instruction does not work. Educational Psychologist 41(2), 75–86 (2006)
6. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. Biometrics 33(1), 159–174 (1977)
7. Magerko, B., Wray, R.E., Holt, L.S., Stensrud, B.: Customizing interactive training through individualized content and increased engagement. In: I/ITSEC (2005)
8. Marsella, S.C., Johnson, W.L., LaBore, C.: Interactive pedagogical drama. In: 4th International Conference on Autonomous Agents. pp. 301–308 (2000)
9. Merrill, M.D.: First principles of instruction. ETR&D 50(3), 43–59 (2002)
10. Miao, Y., Hoppe, U., Pinkwart, P.: Situation creator: A pedagogical agent creating learning opportunities. In: 13th International Conference on AIED. pp. 614–617 (2007)
11. Murray, T., Arroyo, I.: Toward measuring and maintaining the zone of proximal development in adaptive instructional systems. In: 6th International Conference on ITS. pp. 749–758 (2002)
12. Oser, R.L.: A structured approach for scenario-based training. In: 43d HFES Annual Meeting. pp. 1138–1142 (1999)
13. Rieber, L.P.: Seriously considering play. ETR&D 44(2), 43–58 (1996)
14. Riedl, M.O., Stern, A., Dini, D., Alderman, J.: Dynamic experience management in virtual worlds for entertainment, education, and training. ITSSA, Special Issue on Agent Based Systems for Human Learning 4(2), 23–42 (2008)
15. Vygotsky, L.S.: Mind in Society: the Development of Higher Psychological Processes. Harvard University Press, Cambridge, MA (1978)
16. Westra, J., van Hasselt, H., Dignum, F., Dignum, V.: Adaptive serious games using agent organizations. Agents for Games and Simulations pp. 206–220 (2009)