Designing User-Character Dialog in Interactive Narratives: An Exploratory Experiment

Birgit Endrass, Christoph Klimmt, Gregor Mehlmann, Elisabeth André, and Christian Roth

Abstract—**Through interaction with the virtual environment and virtual characters, users are able to influence the storyline of many games. The design choice for the style of interactivity can thereby have a crucial influence on the user's experience. However, only a few approaches evaluate different interaction modalities for one system to investigate the impact of design choice on the users' experience. In this paper, we present an experimental approach in which we first reflect on design alternatives concerning a specific element of interactive narratives—user-character dialog—and then investigate user responses to different design options (round-based dialog versus continuous dialog). Results of an experimental evaluation study show that users tend to prefer continuous interaction in a soap-opera-like game environment using typed text input to communicate with virtual characters that act and react using speech output, although the recognition rate of user utterances of the continuous version was slightly worse compared to the round-based version.**

Index Terms—**Evaluation, game design, interactive drama.**

I. Introduction

INTERACTIVE digital narratives are envisioned and designed to facilitate positive, enjoyable, and moving experiences in their users [1]. In most cases, system creators pursue their own, intuitive, or explicit vision of what they believe to result in a compelling user experience. We believe that formative research that considers user reactions to an interactive narrative could lead to inspiring insights for creators that help optimize the user experience. User studies can therefore help system creators make better choices when several design options are available [2].

In most commercial interactive narratives such as Heavy Rain, user interaction is promoted at certain points of the story only where the user can take decisions by, for example, pressing a button. Approaches that provide continuous interaction, where the user can interact whenever he or she wishes, are still rare. An example includes the interactive drama *Façade*

[3], in which the user can interact via dialog with the virtual characters at any time, providing an experience to the user that resembles an improvisational theater. However, in their follow-up social game *Prom Week* [4], McCoy *et al.* use round-based interaction between the scenes. This is probably due to the extensive calculations that have to be done for the social simulation AI that underlies *Prom Week*. We are very much inspired by their work and aim to find out more about the influence of different interaction styles on user reactions.

From past projects in interactive storytelling, various options for managing user-character dialog are available. In this paper, we investigate the impact of different modes of dialog interaction—round-based versus continuous interaction—on user experience. This way, we contribute to understanding in more detail the links between technological design options and user responses that drive compelling user experiences. The conceivable modes of managing user dialog come with specific advantages and caveats both from a designer's and user's perspective. Exploring potential differences in user reactions to manifestations of different design options regarding dialog-based interaction can help system creators make an informed decision about which kind of option to select. For example, if one mode of managing dialog is particularly costly from a designer's perspective, exploring user reactions to an interactive story prototype that employs this mode will uncover whether gains in user experiences (enjoyable, engaging) can be expected that justify the effort of the dialog mode in question. Various modes of dialog management can be implemented with contemporary technology in interactive storytelling. Designers can make specific choices from a range of options as to the degree of freedom permitted to users as they participate in ongoing conversations among story characters.

In Section II, we present an overview of dialog and interactivity in interactive narratives, pointing out possible advantages or disadvantages of the different design choices. In Section III, we introduce an interactive demonstrator that provides dialog-based interaction in two different interactive modes: round based versus continuous. In Section IV, we present an exploratory user study conducted with 42 participants that interacted with one of the two versions of the system. With the study we investigate the impact of the different dialog modes on the experiences and evaluations of end users. In Section V, we draw conclusions on dialog management in interactive stories and the relationship between design options and user preferences in interactive dialog.

II. Dialog in Interactive Narratives

In many interactive games, the user is limited to reacting to what happens in the environment. In some cases, however, it

is desirable for the user to take the initiative rather than remain limited to reaction only. This interactive dilemma describing the conflict between the author's determinism and the user's degree of freedom has been widely addressed in the area of interactive storytelling (e.g., [5]). Regarding interactive dialog, so far the question of how much user initiative is appropriate has been mainly addressed in the area of natural language processing. Studies conducted by Walker *et al.* [6] on a natural language dialog system seem to indicate that users prefer a system-initiative interface over a mixed-initiative interface. However, robust mixed-initiative interfaces are hard to realize technically, and thus lower user satisfaction with this mixed-initiative interface might also have been caused by the poor performance of the speech recognizer.

Studies on the appropriate level of user initiative in interactive storytelling are still rare. An exception is the work done by Sali *et al.* [7] who investigated three different dialog approaches for game and interactive story telling interfaces. They found that users prefer a natural language interface over those that allow users to select sentences or that make use of an abstract response menu. However, some users had problems with the natural language interface because they found it difficult to figure out what to say in a particular situation.

Mehta *et al.* [8] evaluated design strategies in the interactive drama *Façade* by analyzing the user's experience during conversation breakdowns. Their results indicate that user engagement may be maintained in such situations by providing users with sufficient narrative cues to integrate the virtual characters' reactions into the story. Even during conversational breakdowns, a believable character performance may still provide an enjoyable experience.

The same system was evaluated by Roth *et al.* [9] in comparison to the adventure game *Fahrenheit*, while user experience was investigated either after interaction or after watching a video of the application. Their results indicate that for interactive storytelling systems, such as *Façade*, the user's experience is strongly related to interactivity. This strengthens our idea of evaluating different interactive modes to investigate the impact on the user's experience.

In order to achieve a compelling (entertainment) experience, interactive narratives need to render users' activities both positive in the sense of usability, and meaningful in the sense of comprehensive connections to story development and character behavior [10]. To investigate the impact of different dialog-based interaction styles on the user's experience, we explore two different modes of dialog-based interaction allowing different degrees of autonomy. One version offers a continuous (mixed initiative) dialog interaction to users, whereas the other version offers round-based (system-initiative) dialog possibilities. Empirical research in media psychology suggests that both options will come with specific advantages and problems for an entertaining, meaningful user experience, which will be explored in Sections II-A and II-B.

A. Round-Based Dialog

Round-based dialog limits users' freedom and forces them to become active only on occasions during character conversations that the system enables. From a user perspective, such round-based dialog systems emphasize the autonomy of story characters and force users to pay close attention to ongoing talks, as they need to notice when it is their turn. By limiting the users' freedom in affecting the conversation content and timing, story authors maintain more capacity to elaborate on the autonomous personality of virtual characters and a coherent story plot. The price for this author freedom is that the user's autonomy is reduced to the role of a witness of ongoing conversations permitted to speak only if others (the story characters) allow her/him to do so. While the resulting overall story may be more coherent, the actual user–system interaction may feel less than interactive from a user perspective. This reduction of autonomy may be a disadvantage in terms of entertainment experience, since research into video games has explicitly shown that feelings of autonomy and perceived impact on ongoing game events are key drivers of game enjoyment [11], [12]. On the other hand, round-based dialog comes with the advantage of comprehensiveness and good (perceived) usability. Users can clearly recognize when they can become active; states of uncertainty about when and how to act are less likely to occur if the storytelling system takes the lead in defining phases when interaction is allowed. Since usability and clarity are key preconditions for entertainment experiences (e.g., [13]), the limitation of user autonomy may be compensated for by a gain in usability and the absence of negative impressions of system usability.

B. Continuous Dialog

A design alternative to implementing round-based dialog is to offer users the possibility of continuous interaction. Here, the conversation among story characters is handled in such a way that it allows the users to make a statement (e.g., to type an utterance on the keyboard and submit it to the system) at any time, even if this seems to interrupt a talking story character. While this mode of user-character dialog mirrors a real-life conversation to a greater extent than a round-based dialog, the continuous dialog option comes with greater technical challenges: such as the necessity to make story characters adjust dynamically to new user-generated content. The greater autonomy for users to decide when to become active puts a greater pressure on authors and system designers to be prepared for unconventional, impolite, or incomprehensible behavior displayed by the user. Since the users may intervene at any time, it is more likely that they interact in a way that has not been anticipated by the system designers. Maintaining a coherent story is thus more difficult. From the user experience perspective, however, the greater autonomy should be appreciated, as it allows for a more positive self-experience [14]. On the other hand, the possibility of participating any time in ongoing conversations also increases the pressure on the user to decide when to become active and how to shape the conversation. This may be accompanied by more frequent perceptions of uncertainty and dissatisfaction with how the system responds to one's input.

III. THE STORYTELLING SYSTEM

Our interactive digital storytelling system aims at the creation of a dramatic experience by offering a nonlinear selection of dialog situations in which the user is able to influence the progress and the outcome of the story through his or her interactions and contributions. In order to test the impact of different

Fig. 1. Screenshot of the Virtual Beergarden.

dialog-based interaction styles on the users' experience, we created a user interface that accepts typed text input and a narrative engine that is able to provide appropriate reactions to the user input.

Several software components were implemented: 1) a natural language understanding module that parses the users' utterances into abstract dialog acts; 2) a dialog model and interpreter that controls the narrative structure of the story as well as the dialog and interaction management; and 3) a graphical user interface containing the user input field and the virtual scenario.

A. Virtual Scenario and Story Line

Our scenario is graphically realized with the Virtual Beergarden application [15], shown in Fig. 1. The scenario is located in a Bavarian beer garden containing several groups of nonplaying characters and a user avatar. The characters can freely move through the scene and interact with one another using verbal and nonverbal behavior. For verbal behavior, a text-to-speech engine with different voices is provided, while for nonverbal behavior, over 70 different animations can be performed by each character and customized according to different levels of expressivity. Autonomous low-level behavior, such as positioning toward each other or eye-gaze in conversational settings, is automatically generated by the application.

For our interactive narrative, we chose to create a soap-opera-like story, as this genre is rather popular for users of storytelling systems. The user, represented by a user avatar [Fig. 1(1)], meets a group of female visitors [Fig. 1(2)], a group of male visitors [Fig. 1(3)], and a waitress [Fig. 1(4)]. The user can observe or join these focus groups. During the flow of the story, the user will learn that there is a love story secretly going on, and can actively advise the characters in order to influence the progress and outcome of the story. Dependent on the user's interactions, the characters will reveal their love, ask for help, or follow the user's advice. Different scenario endings are modeled, so that the user can help create a romantic relationship between two characters of his or her choice.

B. Story and Dialog Management

The characters' verbal and nonverbal dialog behavior, as well as the interaction and story management, is realized with the authoring tool SceneMaker [16], [17].

1) Verbal and Nonverbal Behavior: With this tool, dialog content is organized in a set of scenes that are specified in a multimodal scenescript. A scenescript essentially resembles a movie script containing the textual specification of the characters' verbal and nonverbal dialog behavior [Fig. 2(C)]. A scene may contain the spoken utterances of the virtual characters enriched with commands for nonverbal behavior, such as gestures, postures, and facial expressions. Additionally, a scene may contain place holders for variable scene content that is handed over to the scene whenever it is played back.

2) The Logic of Dialog and Story: The dialog logic, the interaction management, and the structure of the story are modeled with a sceneflow [Fig. 2(B)]. A sceneflow is a hierarchical and concurrent state chart which specifies the logical and temporal order in which scenes from the scenescript are executed. The basic building blocks of sceneflows are different types of nodes and edges.

Nodes are organized hierarchically and represent different dialog contexts or stages of a dialog. They contain small executable programs written in a domain-specific scripting language. In particular, these programs contain scene playback commands. These commands are used to generate the virtual characters' behavior during a dialog section.

Edges are transitions between nodes in the sceneflow and are usually guarded by temporal constraints or conditional expressions. A transition is taken if its guarding constraint is satisfied, for example, if the user has made a certain input. Taking an edge can be understood as the transition into a new dialog context or a new story branch.

An author can create an edge including an interruption policy. An interrupting edge is taken as soon as its guarding constraint is satisfied. In this case, any running scene or command of the source node is immediately interrupted. These edges are used whenever a prompt reaction to the user's input is required, for example, if the user wants to interrupt a virtual character's speech or is entering a new focus group. A noninterrupting edge is taken when its guarding constraint is satisfied and the program of the source node has already terminated. These edges allow the modeling of contemporary reactions to the user's input in which the virtual characters finish their ongoing actions first.

With the help of sceneflows, an author is able to model contemporary and adequate reactions to user interactions in consideration of the dialog and story context. Thus, sceneflows can be used to create highly interactive and coherent multiparty dialogs. The effort to create and adopt a dialog and story model is effectively reduced due to the hierarchical and concurrent modeling approach with sceneflows.

3) Example Dialog Flow: In our component-based system architecture, the sceneflow interpreter communicates with the virtual scenario as well as with the user input component of the user interface. Fig. 2 shows the control flow between the system components during an exemplary dialog situation and is explained in the following paragraph.

The user can use the mouse to navigate his or her avatar through the scenario, and approach, for example, the female conversation group in the scene [Fig. 2(1)]. As soon as the user joins the conversation in the female focus group, the Virtual Beergarden application [Fig. 2(A)] signals that event to the
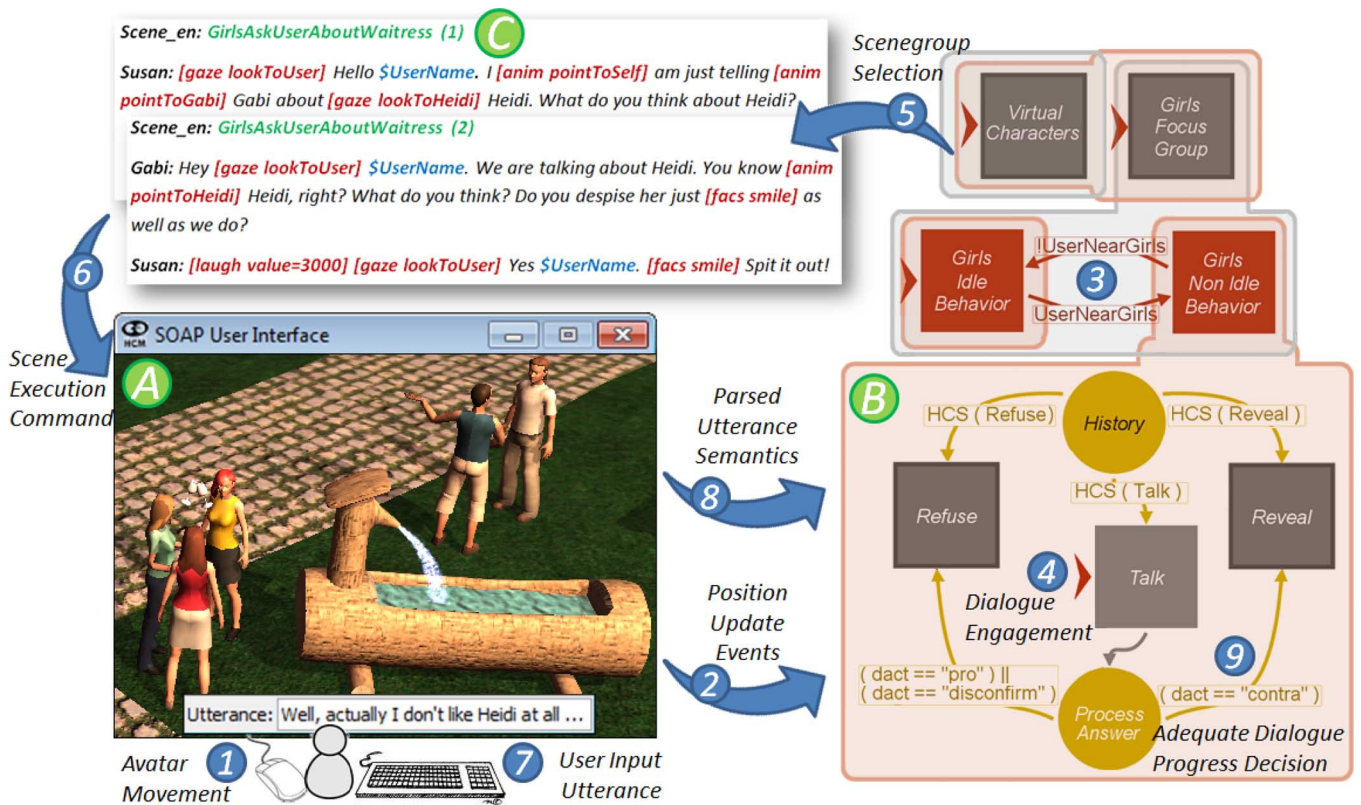
Fig. 2. Control flow between (A) user interface, (B) sceneflow interpreter, and (C) sceneplayer.

sceneflow interpreter [Fig. 2(2)]. The sceneflow, modeled by the author, reacts to this event, first by entering a new dialog context [Fig. 2(3)] and, second, with the start of a new conversation or the resumption of a previously interrupted dialog [Fig. 2(4)]. Technically, this is realized by the transition into a sceneflow node that models the appropriate dialog situation and the decision on how to engage the dialog based on the information in the history memory of this sceneflow node. The substates within the respective sceneflow node contain scene playback commands that are detected by the sceneflow interpreter which sends the request to play such a scene to the sceneplayer [Fig. 2(C)]. In this example, where the user joins the group for the first time, he or she is welcomed by the group with a welcome scene [Fig. 2(5)]. The scenescript contains various variations of that welcome scene and the sceneplayer selects one of these variations based on a selection strategy that has been specified beforehand. The scene is then translated into commands for the Virtual Beergarden application and the verbal and nonverbal actions of the characters are visualized in the scenario [Fig. 2(6)]. Afterwards, the user has the opportunity to respond to the greeting of the characters by typing an utterance into the user interface [Fig. 2(7)]. The user's utterance is translated with a semantic parser into an abstract dialog act and sent to the sceneflow interpreter [Fig. 2(8)]. The sceneflow reacts to the user's dialog contributions with the transition into a certain sceneflow branch that models the appropriate reaction to this specific dialog act [Fig. 2(9)]. The states of the respective sceneflow branch also contain scene playback commands and, thus, the narrative proceeds.

Analogous to the example control flow above, each of the focus groups in our scenario, as well as the user avatar, is mod-

eled as an individual process in a parallel automaton of the sceneflow. In the current version of the system, up to three different dialogs were scripted for each state in order to provide a wider variability for the user, and each script holds verbal as well as nonverbal behavior that could be displayed by the virtual characters. In order to save resources, the scenario is modeled in a manner that dialogs only take place if the user is near a focus group; otherwise, the characters talk gibberish and show random nonverbal communicative behaviors. The automaton modeled for the male characters group includes subautomata that hold the different outcomes of the story.

C. Interaction Modes

To investigate the impact of different dialog-based interaction styles on the user's experience, we implemented two different versions of the interactive narrative: round-based interaction versus continuous interaction. The flow of the overall story is kept as constant as possible for the two versions, to avoid the influence of other factors on the user's experience.

In [18], for example, the importance of the sequence of actions is pointed out by introducing a drama manger that personalizes an interactive story to better match the users' preferences. In line with their findings, Thue *et al.* [19] found out that the sequence of events occurring in an interactive story have a great impact on the users' perceived agency and fun.

Thus, in the prototype used for our evaluation, unsuccessful endings of the narrative (not being able to make a couple) are avoided. Most stages are reachable in both versions and are, therefore, exactly the same in all of them. Other stages, however, can only be reached by the continuous version, since the graphical user interface for the user's typed text input is dis-

abled during agent conversations. For these stages, the characters' reactions to the user's utterances are kept on a similar level of friendliness or politeness as those stages that can be reached by both versions. Nonverbal behavior is also shown in similar amounts and levels of expressivity.

1) Round-Based Interaction: In the round-based interaction style of our demonstrator, dialogs that are held by system-controlled characters are not interruptible. Only when the end of a dialog is reached can the user state an opinion or advise a character, and thus make a contribution to the narrative flow. Therefore, clear interaction prompts are provided by the characters. This interaction style is a lot simpler to model than the continuous version, since only user interactions that are typical for the particular branch are to be expected, and with it the domain for language parsing is limited. If the user is, for example, explicitly asked for his or her opinion on the waitress, it is highly likely that the user will state an opinion rather than try to comment on other things.

2) Continuous Interaction: In this interactive version, the user is able to interact at any time. To this end, the graphical user interface for the user's typed text input is always enabled. The user avatar speaks out the typed utterance while the system-controlled characters of the addressed target group interrupt their current dialog.

This version of our interactive demonstrator is a lot more complex, since agent dialogs need to be interruptible in a manner that characters are able to react to user interaction at any time. Therefore, a much larger set of interaction possibilities needs to be provided as users could interact differently during an ongoing dialog than they might interact after being asked a specific question.

An important issue arising due to this interaction style is the timing of the user's input and the system's output. In case of a user input, the dialog that was currently active is stopped at the utterance level and all descendant nodes may not take any other edges, or execute any further commands. For the correct resumption of an interrupted dialog, a runtime history makes it possible to remember the last substates of the conversational flow.

3) Example Dialogs: The two different interaction strategies as they can occur in our demonstrator are exemplified in Tables I and II. Note that the example of a system-initiated dialog could just as easily occur in the continuous version, should the user not use the possibility to interact proactively, but wait to be addressed by one of the characters instead. In contrast, the dialog exemplifying a continuous dialog cannot occur in the system-initiated version as the user interface is not enabled at that point in time.

D. Natural Language Processing

Providing free text input has certain advantages and disadvantages. On the one hand, free input can encourage the creativity of the user. On the other hand, free input yields many sources of recognition errors, such as spelling errors or misinterpretations. A keyword-based text input is much less error prone. Even more robust performance can be reached with a menu-based input. However, such interaction styles are easily perceived as boring, less intuitive, and restrictive.

TABLE I
EXAMPLE OF THE SYSTEM-INITIATED DIALOG

actor	utterance
Mike	I did not know that Heidi is working in here.
Tom	Me neither. Well I think it is a good choice.
Mike	Yeah. And look at her waitress outfit. I think it is very cute.
Tom	Yes, she definitely looks gorgeous in that skirt with the apron.
Mike	How about you (user name)? What do you think about Heidi?
User (reactive)	Oh I think she is very cute, too!
Mike	So you like her too? I have to admit that I like her very much. But she does not know about my feelings. What do you think I can do?
User (reactive)	You could buy her a bouquet of flowers.
Mike	Great idea, I will give it a try.

TABLE II
EXAMPLE OF THE CONTINUOUS DIALOG

actor	utterance
Mike	I did not know that Heidi is working in here.
Tom	Me neither. Well I think it is a good choice.
Mike	Yeah. And look at her waitress outfit. I think it is very cute.
User (proactive)	True. She looks very cute in it.
Mike	So you like her too? I have to admit that I like her very much. But she does not know about my feelings.
User (proactive)	You could just tell her how you feel.
Mike	Great idea, I will give it a try.

In order to integrate natural language understanding into our interactive scenario, we use the Semantic Parser SPIN [20]. This parser is especially suitable because it enables the analysis of free word order languages and because it copes well with faulty or incomplete text input. We think this is of particular importance when developing an interactive narrative, as the user's input is mainly spontaneous, not necessarily well formulated, and sometimes grammatically incorrect. For that reason, a spell checker was integrated into this component.

The parser analyzes the user's input and maps it to abstract dialog utterances which are sent to the behavior modeling tool for further processing. In order to parse the user's input, rules to the dialog utterances were defined as well as a lexicon that holds words frequently occurring in the domain under consideration of syntactic and semantic categories.

Different speech acts were created for our scenario, each holding a wide variety of possible sentences to be parsed into the speech acts. These speech acts can either be story specific or of a general nature that can be reused for different stories.
- Story specific: guy likes you, girl likes you, guy likes waitress, girls talk about you, I tell waitress, you talk to waitress, buy gift, invite, kiss, give up, work, order.
- General: greeting, farewell, confirm, disconfirm, pro, contra, compliment, insult, don't know, boring, yes, no.

Rules for the semantic parser were refined after a test run of the system with five users, taking into account their typed

utterances. In that manner, we were able to get a clearer insight into what potential users might say to the virtual characters.

IV. EXPLORATORY USER STUDY

An experimental user study was conducted to explore whether the substantial difference that deciding between continuous or round-based user dialog would make from a designer's point of view is also reflected in the end-user experience. The research question was, therefore, whether the user experience would be different in an interactive storytelling prototype that involves continuous dialog compared to one that involves round-based dialog.

A. Study Design and Background

Conceptually, the term "user experience" was grounded on previous theoretical and empirical work [1], [2]. Based on theory and research in media psychology, a dimensional framework of possible user responses to interactive stories was applied. It included diverse modes of experience, such as suspense, curiosity, flow, and effectance (perceived causal influence on the story), as well as important preconditions of meaningful experience, such as perceived usability or character believability. For this conceptual framework, a self-report measurement tool is available [2] and was taken as a basis in the present study. We implemented a self-report measurement tool that was expanded [2] by the addition of the experiential aspect of autonomy [14], which is particularly relevant to the comparison of continuous versus round-based dialog: The former might cause stronger perceptions of autonomy, whereas the latter may make users perceive constrained autonomy, as users need to wait until it is their turn to speak during interaction with the story characters.

Moreover, participants were asked about which "conventional" type of (media) experience the interactive story had reminded them. These questions were used to determine whether shifting between dialog modes affects users' schematic perception of the interactive story (e.g., if the system was experienced like a video game or an improvisation theater).

In total, 42 university students (mean age: 22 years, 30 females) participated in the study. They were randomly assigned to use either the interactive narrative with continuous $(n = 20)$ or round-based dialog $(n = 22)$. After receiving a brief introduction to the system, their exposure typically lasted for five to ten minutes. Afterwards, participants filled in a questionnaire about the various dimensions of the user experience and then received a financial compensation.

For all scales of the users' experiences and the items about which conventional media experience was perceived to be similar, index or item means were computed and compared between participants exposed to either the continuous or the round-based dialog. T-test statistics were applied to determine group differences of particular importance.

B. Results and Discussion

In total, the 42 users of our evaluation study performed 403 dialog acts and interacted for more than four hours with the system. Table III summarizes the users' interactions with the two different dialog modes. Although not statistically significant, users of the continuous version interacted with the char-

TABLE III
AVERAGE NUMBER OF TURNS AND LENGTH OF INTERACTION AS WELL AS RECOGNITION RATE COMPARING THE TWO INTERACTIVE VERSIONS

	round-based	continuous	p (t-test)
average turns	7.86	11.5	0.15
average length	6:42	5:23	0.21
recognition rate	49 %	40%	0.29

TABLE IV
SELECTED FINDINGS ON USER EXPERIENCE. SCALES USED FIVE-POINT RATINGS RANGING FROM 1 (NOT AGREE AT ALL) TO 5 (FULLY AGREE)

	continuous		round-based		t-test
Dimension	M	SD	M	SD	p
Autonomy	2.68	0.93	2.17	0.79	0.055
Curiosity	3.78	0.77	3.33	0.92	0.095
Suspense	3.16	0.74	2.82	0.79	n.s
Enjoyment	3.45	0.95	3.07	0.94	n.s

acters more than in the round-based version. Thus, as expected, users interacted at points in the storyline where it was not intended in the round-based version.

In line with Walker *et al.*'s findings [6], users interacted longer using the round-based version. Although we did not intend for the user to solve a task, the interactive scenario ended with a "happy ending" when the user successfully created a virtual couple. Thus, users of the continuous version were able to reach their "goal" faster than users of the round-based version. Interactions in the continuous version ranged from 2:17 to 16:41 min. Very short interactions in the continuous version occurred in cases where the user unexpectedly triggered an ending event at an early stage of the storyline, for example, by suggesting a dinner invitation without yet being asked for advice.

Although, in most measured experiential dimensions, user ratings did not differ between the two system versions, some informative group differences did emerge. Most importantly, users who interacted with the system through continuous dialog reported greater perceived autonomy and curiosity and were less disappointed in terms of engagement (see Table IV).

Results on the questions about the scenario perception revealed that users rated the experience as more similar to improvisation theater than the users that were confronted with round-based dialog (see Table V). The continuous version of our system was designed to resemble a virtual improvisational theater. Our evaluation study, however, reveals that our story reminds users rather of playing a video game in both cases. This perception does not necessarily need to stand in the light of an improvisational theater experience, since graphics and virtual worlds are directly adopted from video gaming. Nevertheless, our findings indicate that a continuous interactive dialog style enhanced the user's perception toward an improvisational theater experience, which we consider a step in the right direction. Interestingly, the continuous version was perceived as more similar to watching a movie than the round-based version. We believe this is due to the fact that, in the round-based version, the user is interrupted from watching the characters and forced to interact at certain points of the story, while, in the continuous version, the user can freely interact during ongoing conversations and thus might not feel interrupted from watching the story.

TABLE V
Selected Findings on User Expectations and Scenario Perception. Scales Used Five-Point
Ratings Ranging From 1 (Not Agree at All) to 5 (Fully Agree)

	continuous M	continuous SD	round-based M	round-based SD	t-test p
expectation					
"I had the expectation that the experience would be more engaging"	1.75	0.97	2.18	0.59	0.09
perception					
"The experience reminded me of playing a video game"	4.25	1.33	4.82	0.39	0.06
"The experience reminded me of watching a movie"	2.45	0.99	1.77	1.15	0.049
"The experience reminded me a little of playing improvisation theater"	2.10	1.33	1.36	0.49	0.02

Unlike Walker *et al.* [6], we did not observe a user preference for the constrained interface. In contrast to them, we did not use a speech-based interface but presented the users with a robust language interpretation component, by making use of a text-based interface with a spelling checker. In addition, we did not focus on a task-oriented system, such as the spoken language interface to the e-mail system investigated by Walker *et al.* [6]. In their study, the system-initiated version was preferred due to the fact that the interface was easier for the user to learn and more predictable. However, we provide a storytelling scenario that is supposed to serve for entertainment purposes. As a result, solving the task efficiently and having a predictable and simple interface were not anticipated to be equally important to the user. The aim of our system (entertainment) could thus have influenced the users' preference toward the continuous version.

V. Conclusion and Future Work

Our findings indicate that shifting between technologically different options of dialog management did not affect user experience in our interactive environment in a fundamental way. Many conceptually relevant dimensions of the user experience were found to be equal in both experimental groups, and only a few (nearly) significant differences emerged in our study. These differences are highly interesting, however. They suggest that users who interact via the typed text input with characters communicating via spoken language value continuous dialog. This comes with greater technological requirements, in terms of greater autonomy and greater curiosity about how the story will evolve. While the recognition rate of user utterances was rather low in both conditions, users showed a preference for the continuous version, even though the performance was still worse than in the round-based version. Lower system recognition rates thus did not cause weaker user preference, as the experiential advantages of continuous dialog management overruled the greater functionality problems. In our future work, we will aim at improving our system regarding the recognition of user utterances using interactions that occurred during the current evaluation study. Improving the recognition rate hopefully will enhance user experience for both versions and might give some further insight into the user's preference.

Moreover, users' comparisons of our interactive prototype with previous media experiences shifted when the dialog mode was changed: continuous dialog was perceived to be closer to film and improvisation theater experiences, whereas users judged round-based dialog to be more similar to video game play, probably to classic menu-based adventure games. So overall, the technologically more ambitious design option of continuous dialog seemed to contribute to a more unique, novel kind of user experience in our interactive soap-opera-like game environment, whereas the less demanding option of round-based dialog directed the users' perceptions toward well-known experiences of interactive entertainment.

Our descriptive findings indicate, however, that changes in dialog management altered the experience for users of our system only slightly, whereas some aspects of user judgment emerged as stable across all conditions. Specifically, for users of both system versions, the impression of similarity to playing a video game was very strong. This finding certainly originates from the use of game-like graphics and a playful narrative setting. But it also shows that dialog management is only one element that facilitates the user experience. Depending on the design vision of an interactive narrative (e.g., to depart from the familiar video games' look and feel) and the target audience (e.g., literature students instead of gamers), more radical changes to the current system than only a shift in dialog management would be required in order to end up with the desired overall user experience. In turn, the current study demonstrates that dialog management as an isolated system feature can make a difference in user responses in its own right, as observed through interaction with our prototype. So our results suggest that in our dialog-based interactive story, the choice of interaction design can help provide more perceived autonomy and more of an improvisation theater kind of experience to users, compared to an experience similar to playing a video game. Of course, standardized measures do not tell the full story; qualitative approaches with single users are equally important. Yet the fact that theory-based, standardized measures (such as [2]) reveal interpretable and relevant effects of design decisions clearly indicates that quantitative-experimental approaches can make important contributions during system development.

The observation of rather small group differences in our study needs some methodological reflection. First, user ratings were obtained from a prototype system that provides a rather short and simple narrative. The fact that, within such a brief and not very complex context, dialog design made a small yet measurable empirical difference suggests that, with a full interactive narrative, the impact of dialog options on user experience would be much stronger and profoundly relevant to the overall user experience. We thus aim at a new iteration of the system by adding more variations to the characters' behavior and a longer narrative to provide a larger and longer playable experience. Second, the present study has only compared manifestations of different design options that are related to one specific question, namely dialog management. In synthesizing interactive storytelling systems, many more of these design decisions need to be taken. If system creators could make each of the most important of these decisions based on empirical results like those obtained from the current study, they could harmonize the system components

toward a truly optimized end-user experience. The relevance of empirical findings on one particular design option thus could be multiplied across the many issues that system inventors have to consider in interactive storytelling, as well as across manifold different game designs.

REFERENCES

[1] C. Klimmt, C. Roth, I. Vermeulen, P. Vorderer, and F.-S. Roth, "Forecasting the experience of future entertainment technology: Interactive storytelling and media enjoyment," *Games Culture, J. Interactive Media*, vol. 7, no. 3, pp. 187–208, 2012.

[2] I. Vermeulen, C. Roth, P. Vorderer, and C. Klimmt, "Measuring user responses to interactive stories: Towards a standardized assessment tool," in *3rd International Conference on Interactive Digital Storytelling (ICIDS 2010)*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2010, vol. 6432, pp. 38–43.

[3] M. Mateas and A. Stern, "Towards integrating plot and character for interactive drama," in *Proc. Working Notes Soc. Intell. Agents: The Human in the Loop Symp.*, 2002, pp. 113–118.

[4] J. McCoy, M. Treanor, B. Samuel, B. Tearse, M. Mateas, and N. Wardrip-Fruin, "Authoring game-based interactive narrative using social games and Comme il Faut," in *Proc. 4th Int. Conf. Festival Electron. Literature*, 2010 [Online]. Available: https://games.soe.ucsc.edu/sites/default/files/TheProm-ELOAI.pdf

[5] F. Peinado and P. Gervas, "Transferring game mastering laws to interactive digital storytelling," in *Technologies for Interactive Digital Storytelling and Entertainment*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2004, vol. 3105, pp. 48–54.

[6] M. A. Walker, J. Fromer, G. Di Fabbrizio, C. Mestel, and D. Hindle, "What can I say?: Evaluating a spoken language interface to email," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, 1998, pp. 582–589.

[7] S. Sali, N. Wardrip-Fruin, S. Dow, M. Mateas, S. Kurniawan, A. A. Reed, and R. Liu, "Playing with words: from intuition to evaluation of game dialogue interfaces," in *Proc. 5th Int. Conf. Found. Digit. Games*, New York, NY, USA, 2010, pp. 179–186.

[8] M. Mehta, S. Dow, M. Mateas, and B. MacIntyre, "Evaluating a conversation-centered interactive drama," in *Proc. 6th Int. Joint Conf. Autonom. Agents Multiagent Syst.*, 2007, DOI: 10.1145/1329125.1329135.

[9] C. Roth, C. Klimmt, I. Vermeulen, and P. Vorderer, "The experience of interactive storytelling: Comparing Fahrenheit with Façade," in *Entertainment Computing—ICEC 2011*. Berlin, Germany: Springer-Verlag, 2011, vol. 6972, pp. 13–21.

[10] M. Cavazza, J. Lugrin, D. Pizzi, and F. Charles, "Madame Bovary on the holodeck: Immersive interactive storytelling," in *Proc. ACM Multimedia*, 2007, pp. 651–660.

[11] R. Ryan and C. Rigby, "The motivational pull of video games: A self-determination theory approach," *Motiv. Emotion*, vol. 30, pp. 347–363, 2007.

[12] C. Klimmt, T. Hartmann, and A. Frey, "Effectance and control as determinants of video game enjoyment," *CyberPsychol. Behav.*, vol. 10, no. 6, pp. 845–847, 2007.

[13] , M. Blythe, K. Overbeeke, A. Monk, and P. Wright, Eds*., Funology: From Usability to Enjoyment*. Dordrecht, The Netherlands: Kluwer, 2003.

[14] R.-M. Ryan, C.-S. Rigby, and A. Przybylski, "The motivational pull of video games: A self-determination theory approach," *Motiv. Emotion*, no. 30, pp. 347–363, 2006.

[15] I. Damian, B. Endrass, P. Huber, N. Bee, and E. André, "Individualized agent interactions," in *Motion in Games*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2011, vol. 7060, pp. 15–26.

[16] G. Mehlmann, B. Endrass, and E. André, "Modeling parallel state charts for multithreaded multimodal dialogues," in *Proc. 13th Int. Conf. Multimodal Interaction*, New York, NY, USA, 2011, pp. 385–392.

[17] P. Gebhard, G. Mehlmann, and M. Kipp, "Visual scenemaker—A tool for authoring interactive virtual characters," *J. Multimodal User Interfaces*, vol. 6, no. 1–2, pp. 3–11, 2012.

[18] H. Yu and M. Riedl, "A sequential recommendation approach for interactive personalized story generation," in *Proc. 11th Int. Conf. Autonom. Agents Multiagent Syst.*, 2012, pp. 71–78.

[19] D. Thue, V. Bulitko, M. Spetch, and T. Romanuik, "A computational model of perceived agency in video games," in *Proc. 7th AAAI Conf. Artif. Intell. Interactive Digit. Entertain.*, 2011, pp. 91–96.

[20] R. Engel, "Robust and efficient semantic parsing of free word order languages in spoken dialogue systems," in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, 2005, pp. 3461–3464.

Birgit Endrass received the Ph.D. degree (Dr. rer. nat.) in computer science from Augsburg University, Augsburg, Germany, in 2012.

She is a faculty member at Augsburg University. In her research, she focuses on the integration of culture-specific behaviors into the behavior models of virtual agents, and intuitive interaction design.

Christoph Klimmt received a university diploma and the Ph.D. degree in media management from the Hochschule für Musik, Theater und Medien Hannover, Hannover, Germany.

He is a Professor and Director of the Department of Journalism and Communication Research, Hochschule für Musik, Theater und Medien Hannover. Prior to that he was an Assistant Professor at the Department of Communication, University of Mainz, Mainz, Germany. His research interests include media entertainment, video games, processing and effects of media and new media technologies.

Gregor Mehlmann received the B.S. degree in computer science and the M.S. degree (honors) in computer science from Saarland University, Saarbrücken, Germany, in 2007 and 2009, respectively. He is currently working toward the Ph.D. degree in computer science at Augsburg University, Augsburg, Germany.

He is a Researcher and Lecturer at the Lab for Human Centered Multimedia, Department of Computer Science, Augsburg University. Prior to this, he was a Research Associate at the German Research Centre for Artificial Intelligence, Saarbrücken, Germany.

Elisabeth André is a full Professor of Computer Science and Chair of the Lab for Human-Centered Multimedia at Augsburg University, Augsburg, Germany. Her interests are in embodied conversational agents, multimodal interfaces, and social signal processing.

Dr. André is a member of the German Academy of Sciences Leopoldina, the Academy of Europe, and AcademiaNet.

Christian Roth is currently working toward the Ph.D. degree focusing on the enjoyment of interactive storytelling from a media-psychological and communication scientific perspective at the VU University Amsterdam, Amsterdam, The Netherlands.

He is a Researcher at the VU University Amsterdam. Within his research he analyses various dimensions underlying fun and develops a user experience model specific for storytelling applications.