META MATTERS

IN INTERACTIVE
STORYTELLING

AND SERIOUS
GAMING

A PLAY ON
WORLDS

BY

JEROEN LINSSSEN

The meta matters when interacting with stories and (serious) games. This dissertation discusses how the world surrounding such experiences can influence both human users and virtual agents. People can reflect on their experiences and take gained insights into account to improve their interaction and learn from it. Virtual agents can reason about real-world considerations to adapt their behaviour to improve the stories they tell and gaming experience they offer.

Jeroen Linssen is an avid reader and gamer. His background in Cognitive Artificial Intelligence (University of Utrecht) allowed him to research how virtual agents and human users can collaborate in interactive storytelling and serious gaming at the Human Media Interaction Research Group (University of Twente).

ISBN 978-90-365-4343-9

**META MATTERS
IN INTERACTIVE STORYTELLING
AND SERIOUS GAMING**

**A PLAY ON WORLDS
BY
JEROEN LINSSEN**

GRADUATION COMMITTEE

Chairman and Secretary:

Prof.dr. P. M. G. Apers

Promotor:

Prof.dr. D. K. J. Heylen

Assistant-promotor:

Dr. M. Theune

Members:

Prof.dr. R. Aylett

Dr. R. Bidarra

Prof.dr. J. van Hilligersberg

Prof.dr. T. W. C. Huibers

Prof.dr. J-J. Ch. Meyer

Dr. N. Szilas

PARANYMPHS

Dr. M. Bruijnes

Ing. T. F. de Groot

The research reported in this dissertation was carried out at the Human Media Interaction group of the University of Twente.

COMMIT/

The research reported in this dissertation was supported by the Dutch national program COMMIT.

CTIT

CTIT Ph.D. Thesis Series No. 17-432
Centre for Telematics and Information Technology
P.O. Box 217, 7500 AE
Enschede, The Netherlands

SIKS Dissertation Series No. 2017-21
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

ISBN: 978-90-365-4343-9

ISSN: 1381-3617 (CTIT Ph.D. Thesis Series No. 17-432)

DOI: 10.3990/1.9789036543439 (<https://doi.org/10.3990/1.9789036543439>)

© 2017 Jeroen Linssen, Enschede, the Netherlands

META MATTERS IN INTERACTIVE STORYTELLING AND SERIOUS GAMING
(A PLAY ON WORLDS)

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr. T. T. M. Palstra,
on account of the decision of the graduation committee
to be publicly defended
on Thursday, May 4, 2017 at 16:45

by

Johannes Maria Linssen
born on June 11, 1987
in Venray, the Netherlands

This dissertation has been approved by:

Prof.dr. D. K. J. Heylen, University of Twente, NL (supervisor)

Dr. M. Theune, University of Twente, NL (co-supervisor)

SUMMARY (ENGLISH)

Stories, in games, novels, films, and all media alike, help us understand our world. Stories provide an escape from the real world, but they also reflect reality. They provide situations that may previously have been unimaginable, leading us to think about concepts in different ways. We learn through stories, we construct our mental model of the world through them.

In this dissertation, I investigate the use of storytelling techniques in the domains of interactive storytelling and serious gaming. The field of interactive storytelling attempts to create and analyse systems that provide an interactive experience to users: the latter can influence the plot of story as it is being told. Serious gaming is concerned with the use of games for purposes other than only entertainment; in this dissertation, I investigate serious games with educational purposes. In both domains, I inspect how both human users and virtual agents (artificial, autonomous entities) can perceive stories.

In Part I of this dissertation, I present my main thesis: the *meta matters* when interacting with interactive storytelling systems and serious games. All stories and games take place in a context: they take place in virtual worlds embedded within the real world. This concept is captured by the term *magic circle*. Virtual worlds are not separated from the real world; they are interconnected. Moreover, events and perceptions in the real world influence people's behaviour in the virtual world and vice versa. My research revolves around the use of such meta-information to improve experiences with interactive storytelling systems and serious games.

In Part II of this dissertation, I delve into the field of interactive storytelling, specifically focusing on the Virtual Storyteller (VST) and its successor, the Interactive Storyteller (IST). The VST provides users with an emergent narrative: stories are generated based on the interactions between virtual agents which play characters in the story. In the IST, users can take control of one of the characters, substituting a virtual agent, and influence how the story plays out. I first investigated how the meta matters to these systems by analysing interactions between pairs of children and the IST. This analysis showed that children, apart from only enacting *in-character* (IC) behaviour, actively go meta to reason about their plans and discuss their experience amongst each other (*out-of-character* behaviour, OOC). We created a scheme that allows for the annotation of this communication to discern the types of IC and OOC behaviour. With this scheme, we could see how the children's perspectives toward their interaction changed over time.

Using the VST, we investigated how the meta can matter to its virtual agents. The generation of stories in the VST relies on the characters in the story world performing actions that advance the story. However, this does not necessarily entail that characters will take those actions. For example, if they can go about their way and never meet each other, possible conflicts can be avoided. To circumvent this problem, we outfitted the virtual agents with OOC techniques that allow them

to go meta to plan actions that result in conflicts. For example, a character can go meta to assume (ooc) that an object it is looking for is at the location of its antagonist, leading to a confrontation between the two. This way, generated stories can become more interesting.

In Part III of this dissertation, I investigate serious games and take an approach which is similarly structured to that of Part II. We created a digital serious game called LOITER which features a police officer (the player character) who has to convince a small group of loitering juveniles to cease their loitering. The overarching learning objective of this game is that players should improve their social awareness, being able to interpret how their own behaviour influences that of others and vice versa. To bring about this understanding in players, I designed feedback mechanics that are shown during the game and give insight into the thoughts of the virtual juveniles: thought bubbles and flashbacks. These mechanics allow players to briefly go meta while they are playing the game, giving them the opportunity to reflect and learn from their interaction.

Secondly, I investigated how the virtual agents can further support the learning objectives of LOITER. I designed mechanics that allow the virtual agents to make use of meta-information to adapt the game's difficulty to players. Based on a player's competence, the virtual agent can determine whether that player has reached a certain level of social awareness. If so, the virtual agent adapts its behaviour to be less cooperative so that the player will have a harder challenge to convince the virtual juvenile to stop loitering. As this change in behaviour may come across as incoherent, the game presents motivations for such a change in the form of either a thought bubble or flashback. This way, LOITER can provide a difficulty level that is tailored to individual players.

This dissertation concludes with Part IV, in which I summarize the above points and provide an outlook on future work. One such research direction concerns envisioned improvements to the Interactive Storyteller, most notable how such an interactive storytelling system can receive player input. Future work should also expand on the possibilities of virtual reality (VR) for serious gaming. We already took a small step in this direction by implementing LOITER in a multi-modal VR environment. Here, again, the challenge is to correctly interpret players' intentions through the available channels for input.

In the end, this dissertation shows that, by enabling and stimulating both virtual agents and human users to use meta-information, interactive storytelling systems and serious games can be improved.

SAMENVATTING (NEDERLANDS)

Verhalen, in games, romans, films, en alle vergelijkbare media, helpen ons om de wereld te begrijpen. Verhalen bieden de mogelijkheid om te ontsnappen aan de echte wereld, maar reflecteren ook de realiteit. Ze verschaffen situaties die we voorheen niet hadden kunnen voorstellen, zodat we na kunnen denken over concepten op verschillende manieren. We leren dankzij verhalen, dankzij hun creëren we een mentaal model van de wereld.

In dit proefschrift onderzoek ik het gebruik van verhaaltechnieken in de domeinen *interactive storytelling* (interactieve verhalen) en *serious games* (educatieve spellen). Het veld van interactieve verhalen poogt systemen te maken en analyseren om een interactieve beleving aan te bieden aan gebruikers: die laatsten kunnen het plot van het verhaal beïnvloeden terwijl het verteld wordt. *Serious games* zijn spellen met een ander doel dan louter vermaak; in dit proefschrift onderzoek ik *serious games* met educatieve doeleinden. In beide domeinen analyseer ik hoe menselijke gebruikers en virtuele agenten (kunstmatige, autonome entiteiten) verhalen waarnemen.

In Deel I van dit proefschrift presenteer ik mijn hoofdthese: de *meta maakt uit* tijdens het interageren met interactieve verhaalsystemen en educatieve spellen. Alle verhalen en spellen spelen zich af in een context: ze spelen zich af in virtuele werelden die ingebed zijn in de echte wereld. Dit concept is omvat in de term de *magic circle*. Virtuele werelden zijn niet gescheiden van de echte wereld; zij zijn onderling verbonden. Gebeurtenissen en percepties in de echte wereld beïnvloeden bovendien het gedrag van mensen in de virtuele wereld en vice versa. Mijn onderzoek draait om het gebruiken van meta-informatie met als doel belevingen met interactieve verhaalsystemen en educatieve spellen te verbeteren.

In Deel II van dit proefschrift verdiep ik me in het veld van interactieve verhalen, waarbij ik me concentreer op de Virtual Storyteller (VST) en haar opvolger, de Interactive Storyteller (IST). De VST biedt gebruikers een emergente vertelling aan: verhalen worden gegenereerd op basis van de interacties van virtuele agenten die karakters spelen in het verhaal. In de IST kunnen gebruikers de rol van een virtuele agent overnemen en een karakter spelen om op die manier te beïnvloeden hoe dat verhaal verloopt. Ik heb onderzocht wat de meta uitmaakt voor deze systemen door interacties tussen kinderparen en de IST te analyseren. Deze analyse toonde aan dat kinderen, naast het uitdragen van *in-karakter* (IK) gedrag, actief meta gaan om na te denken over hun plannen en hun ervaringen met elkaar te bespreken (*buiten-karakter* gedrag, BK).

Met behulp van de VST onderzochten we hoe de meta uitmaakt voor haar virtuele agenten. Het genereren van verhalen in de VST geschiedt doordat karakters in de verhaalwereld acties uitvoeren die het verhaal laten vorderen. Dit betekent echter niet noodzakelijk dat de karakters daadwerkelijk die acties uitvoeren. Zij kunnen bijvoorbeeld mogelijke conflicten ontwijken, als zij vrijelijk hun gang kunnen gaan en elkaar nooit ontmoeten. Om dit probleem te overbruggen rustten

we de virtuele agenten uit met BK-technieken die hen in staat stellen om meta te gaan zodat ze acties kunnen plannen die uitmonden in conflicten. Een karakter kan bijvoorbeeld (BK) aannemen dat een object waar zij naar zoekt zich op de locatie van een antagonist bevindt, wat leidt tot een confrontatie tussen deze twee karakters. Op deze manier kunnen gegenereerde verhalen interessanter worden.

In Deel III van dit proefschrift onderzoek ik educatieve spellen met eenzelfde benadering als die in Deel II. We hebben een digital, educatief spel gemaakt genaamd LOITER, waarin een politieagent (de speler) een groep hangjongeren moet overtuigen om hun overlastgevende gedrag te staken. Het overkoepelende leerdoel van dit spel is dat spelers hun sociale inzicht moeten verbeteren, dat wil zeggen, in staat zijn te interpreteren hoe hun eigen gedrag dat van anderen beïnvloedt en vice versa. Om dit begrip over te brengen op spelers heb ik feedback-mechanismes ontworpen die tijdens het spel getoond worden om inzicht te geven in de gedachten van de virtuele jongeren: gedachtewolkjes en flashbacks. Deze mechanismes geven spelers de kans om kortstondig meta te gaan terwijl ze het spel spelen, zodat ze de gelegenheid hebben te reflecteren en te leren van hun interactie.

Hierna heb ik bekeken hoe virtuele agenten verder de leerdoelen van LOITER kunnen ondersteunen. Ik heb mechanismes ontworpen die virtuele agenten toestaan om met meta-informatie de moeilijkheidsgraad van het spel aan te passen aan spelers. Gebaseerd op de competentie van een speler kan de virtuele agent bepalen of die speler een bepaald niveau van sociaal inzicht heeft bereikt. Als dat het geval is past de virtuele agent zijn gedrag aan om zich minder coöperatief te gedragen, zodat de speler een grotere uitdaging heeft om de virtuele jongere te overtuigen te stoppen met overlast veroorzaken. Omdat deze verandering in gedrag incoherent kan overkomen toont het spel verklaringen voor zulke veranderingen in de vorm van gedachtewolkjes en flashbacks. Op deze manier kan LOITER een moeilijkheidsniveau aanbieden dat is toegespitst op individuele spelers.

Dit proefschrift sluit met Deel IV, waarin ik de bovenstaande punten samenvat en een blik op de toekomst werp. Eén onderzoeksrichting heeft betrekking op beoogde verbeteringen van de Interactive Storyteller, in het bijzonder hoe een dergelijk interactief verhaalsysteem om kan gaan met spelerskeuzes. Toekomstig onderzoek zal ook de mogelijkheden van virtual reality (VR) moeten exploiteren voor educatieve spellen. We hebben reeds een kleine stap in die richting genomen door LOITER te implementeren in een multimodale VR-omgeving. Ook hier is het een uitdaging om de intenties van spelers correct te interpreteren middels de beschikbare kanalen voor spelerskeuzes.

Tot slot laat dit proefschrift zien dat interactieve verhaalsystemen en educatieve spellen verbeterd kunnen worden door het virtuele agenten en menselijke gebruikers zowel mogelijk te maken als te stimuleren om meta-informatie te gebruiken.

SAMEVATTING (LIMBURGS)

Dur verhaole kunne weej liere en de wèlt snappe. Ien dit proefschrift onderzuuk ik ut gebruik van verhaoltechnieke ien interactieve verhaole en serieuze spelle. Ien interactieve verhaole kunde zelf beslisse hoe ut verhaol aflöpt. Van serieuze spelle kunde wat liere. Ik kiek hoe minse en virtuele pupkes verhaole bekieëke.

Ien Deel I van dit proefschrift zeg ik waor dit allemaol op slut: de *meta makt uut*. Alle verhaole en spelle speule zich af ien un context: ze zien virtueel (nie echt), maar hebbe wel te make mit de echte wèlt. Wat dur buuten ut spel gebeurt het un effect op wat dur ien ut spel kan gebeure en andersum. Mien onderzuuk draoit um ut gebruuken hiervan um de belèvinge mit interactieve verhaole en serieuze spelle te verbèttere.

Ien Deel II bekieëk ik urst de interactieve verhaole die te make hebbe mit de Virtual Storyteller (VST) en Interactive Storyteller (IST). Dun urste gebruukt virtuele pupkes die vur zichzelf dinke om een verhaol te make. Ien de IST kunne minse un pupke ien ut verhaol speule. Ik heb gekeke naor hoe kiender mit de IST verhaole make en hoe ze daorbeej meta gaon. Ien de VST hebbe weej gekeke naor hoe virtuele pupkes ruzie mit elkaor könne zuuke, zodat ut verhaol ienteressanter kan worre.

Ien Deel III kieëk ik naor un serieus spel: LOITER. Hierien mot un agent (de speuler) un groep hangjongens aovertuuge om te kappe mit hun herrie. Speulers motten daorbeej inzien dat hun gedrag dat van andere kan bepaole en andersum. Hiervur heb ik dinge gemakt die speulers laoten stilstaon beej wat ze duun: gedachtewolke en flashbacks die laoten zien hoe un hangjonge aover hun dinkt. Hierna heb ik gekeke hoe die virtuele pupkes ut moeilijker kunne make vur speulers. Dit duun ze dur minder goed same te werke mit speulers die wat verder zien ien ut spel. Speulers hebben dan een grotere uitdaging. Om daorvur te zörge dat speulers nie ien de war rake van die verandering in gedrag van de pupkes, laot ik hun uutlegge waarom ze dat duun mit gedachtewolke en flashbacks.

Dit proefschrift sluit mit Deel IV mit unne samevatting en unne blik op de toekomst. Onderzuuk mot kieëke naor hoe spelle en zulke systeme bèter om kunne gaon mit de keuzes van speulers (beejvurbeeld ien virtual reality).

Om af te sluute: interactieve verhaole en serieuze spelle kunne verbètterd worre dur virtuele pupkes en minselijke gebruuikers te stimuliere meta-informatie te gebruuken.

ACKNOWLEDGEMENTS

It is possible I already had some presentiment of my future.

Severian
The Book of the New Sun
Wolfe (1994)

I have long wondered how I would conclude this tale. I knew I would be writing these acknowledgements, hopefully sufficiently thanking each and every one of you that helped me during the course of my PhD-candidacy. Looking back, your number is greater than I had imagined beforehand.

But I'm getting ahead of myself. In a sense, the beginning of most stories is written at the end, and the end at the beginning. When I started my PhD at the Human Media Interaction research group of the University of Twente, I had some idea of what I should end up with after four years. Socially intelligent agents for serious gaming environments, assisting players (police officers in training) to become more socially aware. The road there took many turns: I investigated the fields of interactive storytelling systems, of serious games (for social skills), of social psychology, of educational psychology, of intelligent agents; I dove into the code of the Virtual and Interactive Storytellers, fiddled around with HTML5 and Javascript for LOITER; I talked to police officers, both tutors and those in the field; I visited many a workshop, course, and conference to stay up-to-date with recent advances in the relevant fields of research; I (co-)supervised plenty of students; and I must have taken more turns still that escape my mind right now. With this list, I don't mean to boast. On the contrary, now, more so than ever, I understand that to conduct science is to be aware that one can never know everything, one can only strive to increase one's knowledge. At the end of these four(ish) years, I have come to understand that I know just enough to speak my mind on certain things: hence, this dissertation. Moreover, I understand that much work remains to create truly socially intelligent agents and successful serious gaming environments.

Ending up at this point in my life, I am extremely grateful to those who helped me along the way. Mariët, I don't know where to start. It is truly mind-boggling to me how you managed to support me in so many ways over the years. Whether it was the fifth revision of a paper you still managed to improve even further, helping me to emphasize what I actually wanted to say, asking the ever-so critical questions (and rightly so!), or just sharing an ever-welcome cup of tea, you helped me tremendously. Without you as a guide, I would still be lost along the way here. Dirk, your smiles meant a lot to me, literally. At times, they were commending me on my work; perhaps more often, they expressed gladness at seeing I began to understand what I actually was (or should be) doing. Still, they meant well and I feel that I progressed a lot during my time at HMI; I hope you feel the same.

While we're talking about HMI: it's a place where we do stuff with things. We talk, play games, have lunch, and occasionally dabble in scientific affairs. For a large part, it's a serious research group, but I'm very lucky to have carried out my PhD research at such a welcoming, cosy place. My colleagues, they're awesome. First off, my room mates are a hilarious couple of opposites and similarities. Merel, your tidiness, dry humour, and love for good games will stick with me for a long time. Cristina, your Italian nature is inspiring and enticing. Never change – but take it easy. Jan, thanks for being a most bodacious bro; I hope we'll share many more laughs in the future. Roelof, stay relaxed; thanks for your enthusiastic discussions about politics (but take it easy!). Daniël, Jered, thanks for being good neighbours, not complaining about the frequent visits our room had, and making *Snoozle* a success thus far. My thanks also go to the Fussball pros, Robby & Gijs, for not making me crawl underneath the table too many times. Randy, Khiet, thanks for the fries. Bob, stay metal. Mannes, we had a good run as a multi-agent system to lecture the similarly named course; thanks for being so down to earth. Rieks, your help at making various pieces of the IUALL project a success is much appreciated. Alice, Charlotte & Lynn: thanks so much for your support; you are the three pillars of HMI. Et al. of HMI: thank you for making everything happen.

Last, but definitely not least, my paranymphs: Merijn & Thomas. When I first set foot in HMI, I was a different person than I am now. The both of you have taught me to not just try things but actually do them. I owe you a lot for making me the person right here, right now, more confident and more optimistic. I mean it when I say it has been an honour to have worked alongside you and I believe that, with your help, I have made this PhD into a success.

In carrying out this research, I am also happy to have worked together with the people of RE-liON. Though our collaboration only makes up a small part of this dissertation, I feel that we have made something worthwhile to look back upon. Likewise, I am grateful to the people from the Dutch National Police: Carma, Siemen, Ron, Bert, thanks for enabling me to carry out my research. Another round of thanks goes to the people at T-Xchange, who have made it possible to actually make LOITER. I am of course also much-indebted to the many students who, in some way or another, helped and inspired me to do the work in this dissertation: Chris, Hans, Marissa, Rens, Riccardo, René, Ineke, Carin, and many others. On a grander scale, the same goes for COMMIT/, the Dutch national program which financed the IUALL project of which my research formed a piece. Thanks for making all of this possible and bringing such a varied bunch of researchers together at many stimulating events! In finishing up this dissertation, the comments from the committee already provided healthy food for thought. I thank you all for agreeing to review, discuss, and criticize my work; I'm sure it turned out the better because of this.

I feel that I should also dedicate one sentence to all the novels, films, games, music, photography etc. that inspired and continue to inspire me to carry out my research; the various quotes opening the chapters of this dissertation should indicate what their impact was on me.

Dan nog een korte dankbetuiging aan mijn vrienden en (uitgebreide) familie. Mijn BPC-kameraden; het is altijd gezellig met jullie. Altijd lachen, om wat dan ook. Ik hoop dat er nog vele MTB-weekenden en LAN-party's volgen.

Pap, mam, zoveel dank voor alles. Jullie zorg betekent heel veel voor me; jullie kennen me als iemand van weinig woorden, maar laat me zeggen dat ik hoop dat jullie net zo trots zijn op mij als ik op jullie. Bedankt voor het ouder-zijn. Sven & Moniek, bedankt voor het jonger-zijn en het anders-zijn. Oma, veel dank voor het mogelijk maken mijn Limburgs te verbreden en verbeteren met alle geweldige Venlose woorden en gezegdes. Dit proefschrift is ook voor opa; ik hoop dat hij trots geweest zou zijn. De familie Jorna-Nijhoff: na al die jaren (het zijn er wat) voel ik me nog steeds welkom. Bedankt daarvoor en ook voor de (te) levendige discussies.

Clichés are often used because they are true; lo, this closing paragraph. Leonie, we met quite a while ago and figured out we had a lot in common. Over the years, I think we helped each other in different ways and learned about many new things together. Many a horseback-ride still has a vivid spot in my memory, as do the treks we planned and undertook in various Nordic countries. Your enthusiastic, passionate spirit remains ever stimulating. I hope we travail on many more trails and conquer many more mountains. Though plenty of years are behind us, I hope many more remain. I feel we are still at the beginning of something incredible. Thank you for being here.

Jeroen Linssen
In a train somewhere, 2017-04-17

CONTENTS

I	INTRODUCTION: SETTING THE STAGE	1
1	THE NARRATIVE OF THIS DISSERTATION	3
1.1	A play on worlds	4
1.2	Research context	6
1.2.1	Interactive digital storytelling	6
1.2.2	Serious games	8
1.3	Going meta: agens reputans	11
1.4	Research questions	13
1.5	Outline of this dissertation	14
II	META MATTERS IN INTERACTIVE STORYTELLING	17
2	INTERACTIVE STORYTELLING	19
2.1	Interactive storytelling systems	19
2.1.1	The strong story approach	19
2.1.2	The strong autonomy approach	22
2.1.3	Hybrid approaches	24
2.2	The Interactive Virtual Storyteller	28
2.2.1	The Virtual Storyteller	28
2.2.2	The Interactive Storyteller	33
3	THE USER PERSPECTIVE ON INTERACTIVE STORYTELLING	37
3.1	Interaction with interactive storytelling systems	38
3.2	Study setup	39
3.3	Coherent co-creation	41
3.3.1	Story components	43
3.3.2	Causal connectedness	45
3.3.3	Children's communication	48
3.3.4	Discussion	51
3.4	Perception of interaction	53
3.4.1	Frames of play	54
3.4.2	The P×R annotation scheme	55
3.4.3	Analysis of children's communication	57
3.4.4	Validity of the P×R scheme	60
3.5	Reflections on this chapter	61
4	AGENTS THAT PERCEIVE AND ASSUME TO STEER STORIES	65
4.1	Perceiving and assuming in interactive storytelling	66
4.2	Limitations of omniscience for story generation	67
4.3	Perceptions in the VST	68
4.3.1	Visibility rules	69
4.3.2	Dealing with perceptions	69
4.3.3	Initial character knowledge	70
4.4	Assumptions in the VST	71

4.4.1	Assumption operators	71
4.4.2	Preconditions	72
4.4.3	Executing assumptions	72
4.4.4	Plot control	73
4.5	Evaluation	74
4.5.1	Expanding the story world	74
4.5.2	Example stories	74
4.6	Reflections on this chapter	77
III	META MATTERS IN SERIOUS GAMING	81
5	SERIOUS GAMES	83
5.1	Serious games for social skills training	83
5.1.1	Tactical language training systems	84
5.1.2	deLearyous	86
5.1.3	STRESS	87
5.1.4	FearNot!	88
5.1.5	Nothing for Dinner	88
5.2	LOITER	88
5.2.1	Modelling interpersonal behaviour	89
5.2.2	Learning objectives for social awareness	93
5.2.3	Game design of LOITER	96
6	FEEDBACK IN SERIOUS GAMES	105
6.1	Theories and implementations of feedback	106
6.1.1	Support through feedback	106
6.1.2	Implementations of feedback in serious games	108
6.2	Meta-techniques in LOITER	109
6.2.1	Live action role play	109
6.2.2	Feedback through meta-techniques	110
6.3	Thought bubbles	112
6.4	Experiment: thought bubbles in LOITER-TB	113
6.4.1	Method	113
6.4.2	Results	116
6.4.3	Discussion	117
6.5	Flashbacks	118
6.6	Experiment: flashbacks in LOITER-FB	119
6.6.1	Method	119
6.6.2	Results	121
6.6.3	Discussion	122
6.7	Reflections on this chapter	122
7	ADAPTIVE SERIOUS GAMES	125
7.1	Adaptive difficulty in serious games	126
7.1.1	Tailoring challenge to competence	126
7.1.2	Implementations of adaptive difficulty in serious games	129
7.2	Adaptive stance profiles of virtual characters	132
7.2.1	Design considerations for adaptation	133
7.2.2	Final design of adaptation	134

7.2.3	Justifications	137
7.2.4	Implementation of LOITER-AP	140
7.3	Experiment: adaptation in LOITER-AP	141
7.3.1	Method	142
7.3.2	Results	144
7.3.3	Discussion	149
7.4	Reflections on this chapter	151
IV	CONCLUSION: REFLECTING ON THE RESEARCH	153
8	LESSONS LEARNT	155
8.1	The meta-narrative of this dissertation	155
8.2	Interactive storytelling gone meta	156
8.2.1	How players go meta to co-create stories	157
8.2.2	How agents go meta to perceive, assume and plot stories	158
8.3	Serious gaming gone meta	159
8.3.1	How to support players to meta-think and reflect	160
8.3.2	How agents go meta to adapt a game's difficulty	161
8.4	The meta-paradigm	163
9	FUTURE WORK	165
9.1	Improving interactions with the Virtual Storyteller	165
9.2	LOITER-VR: Social skill training in virtual reality	167
9.3	The narrative of the future	170
	BIBLIOGRAPHY	173
V	APPENDICES	189
A	INTERACTIONS WITH THE INTERACTIVE STORYTELLER	191
B	ADDITIONAL MATERIAL ON LOITER	193
B.1	LOITER-TB	193
B.2	LOITER-FB	195
B.3	LOITER-AP	196
B.3.1	Pseudocode for profile stance calculation	196
B.3.2	Dutch translations of justifications	197
B.3.3	Explanation of game mechanics in LOITER-AP	198

PART I

INTRODUCTION: SETTING THE STAGE

In this part, I reveal what this dissertation is about.

I introduce the fields of interactive storytelling and serious games, discussing the roles human and virtual agents play in these fields. My main thesis is that, to improve systems in these fields, both human and virtual agents need to *go meta*: they need to reflect about what they are doing to adapt their behaviour. In this part, I pose four research questions which I investigate in Parts II and III.

THE NARRATIVE OF THIS DISSERTATION

1

1

In your Primer you have a resource that will make you highly educated, but it will never make you intelligent. That comes from life. Your life up to this point has given you all of the experience you need to be intelligent, but you have to think about those experiences. If you don't think about them, you'll be psychologically unwell. If you do think about them, you will become not merely educated but intelligent.

Constable Moore
The Diamond Age: Or, the Young Lady's Illustrated Primer
Stephenson (1996)

Two children, a boy and a girl, are standing on opposite sides of a table. Intently, they look at the screen embedded in the table, which shows a landscape in a cartoonish style. Three characters seem to inhabit this small, two-dimensional world: Red, a girl with a red riding hood, her Grandmother, and a big bad Wolf.

Suddenly, the Wolf moves to Red: 'Hello!' The Wolf is controlled by the computer, so the children don't know for sure what he might do.

The boy looks across the table into the girl's eyes: 'You have to give him the cake!'

The girl replies: 'It has to be poisoned first.'

The boy picks up on her plan: '(..) I will do that, alright?'

The girl drags the image of Red to Grandmother's house, making her walk over there. Letting her in, the boy makes Grandmother add a few droplets of deadly poison to Red's cake. In the meantime, the Wolf grows ever more hungry and decides to bother Red and Grandmother again.

The children exchange looks, the girl saying: 'Gee, he eats a lot!'

The Wolf enters Grandmother's house as well and snatches the cake from Red: 'Give it to me!'

Both children laugh in expectation of the drama about to befall the Wolf.

The girl selects an action to make Red laugh as well. The table plays a short audio clip of a little girl laughing.

Then, with a loud chomp, the Wolf sets his teeth into the cake, its last meal.

THE above is a short example of what may happen when children play with an interactive storytelling system. While they instruct their characters to take actions, the children are aware of most of the events happening in the story. They can see where their characters are in the storyworld, what they are doing, but they do not necessarily know the plan of the Wolf, the computer-controlled antagonist. Yet the children also know about more than just the storyworld. They are aware that they are playing this story and they know they can discuss which actions to take in the story without the Wolf listening in on them. More than just being characters inside the story world, the children experience the events and possibilities outside of the story world: they have a view that extends beyond it.

In this dissertation, I discuss the *meta matters* of such interactions with intelligent technology. People may reason about their choices, plot, reflect, or simply be amused when they look at things from a meta-perspective. I assert that this perspective can be used to improve existing systems when implemented in a way that lets people use it to enrich their experiences. Furthermore, intelligent systems themselves may also inspect interactions from a meta-perspective to improve the user's experience.

I carried out my research in the Dutch national program COMMIT/, in the project *Interaction for Universal Access* (IUAL).¹ The aim of this project was to investigate how interactive computer systems can be designed in such a way that all people of today's society can interact with them. This holds especially for groups that are often overlooked, such as people with low literacy, children, and elderly. The key of IUAL lies in designing computer systems to have social intelligence and providing tools to make it possible to design for all people. My part in this project was to investigate how interactive technology can support social skills training using techniques from interactive storytelling. This is reflected in the structure of this dissertation: its two central parts revolve around interactive storytelling and serious games for social skills training.

To properly introduce the basis of my research, I discuss some seminal works about humans as storytellers and players in Section 1.1. I follow up on this by providing a context for my research in Section 1.2. In Section 1.3, I expand upon my idea of people and systems *going meta* to improve their interactions. Section 1.4 provides the research questions I investigate in this dissertation. In Section 1.5, I conclude this introductory chapter with an overview of this dissertation's narrative.

1.1 A PLAY ON WORLDS

Stories help us understand the world we inhabit. By reading articles, news feeds, updates on social media, and conversing with people, we build a mental model of our world. We causally link events and weave our interpretations into stories that we use to make sense of the world. This way, we learn about the universe: our surroundings, social relations, physics, morality, biology, et cetera. Similarly, by reading novels, experiencing plays, watching films and series, we explore

1 See <http://www.commit-nl.nl>.

a multitude of possible situations in alternate realities. Again, by interpreting the events and actions, the metaphors and allusions, and the characters' coping with their emotions in these stories, we learn about our own world. These are the thoughts underlying Ricoeur's concept of a *narrative identity* (Ricoeur, 1983): humans construct their identity through stories, both by telling and listening to them. Niles (1999) coined the term *homo narrans* to describe humans as both creators and interpreters of stories. Such a view is also prevalent in Campbell's writings, which delve into the archetypical structure of a hero's journey (Campbell, 1949). All cultures exist by the grace of their ability to tell stories which help them shape their lives, for example, through a story's morals, the catharsis which a story may provide, or by following rituals.

Fisher (1989) suggests a paradigm of *narrative rationality* through which we can understand how we experience and learn from stories. This paradigm informs us that a narrative can only be useful to the reader² when it satisfies *narrative coherence* and *narrative rationality*. The former requires a story to 'hang together' to make sense, while the latter entails that we only accept a story if it matches our experiences and beliefs. Recent research has shown that, indeed, experiencing and also creating stories helps people to understand situations by determining the causality between events (A. D. Brown, Stacey & Nandhakumar, 2008). Similarly, Graesser, Olde and Klettke (2002) looked at several studies on text comprehension and concluded that 'narrative text is recalled approximately twice as good as expository text and is read approximately twice as fast,' showing the benefit of stories as instructional material.

Rather than being strictly linear, stories can be interacted with in several ways to influence their outcomes. For example, Choose Your Own Adventure (CYAO) novels allow people to make decisions for the characters in the story. Current advances in supporting the creation of stories with technological means also allow for stories to become even more interactive. I describe the field of *interactive storytelling* research in more detail in the following section.

More so than only being readers of stories, people can also create narratives themselves, for example, by role-playing and improvising how a story turns out. At the start of his book *Homo Ludens* (the playing man), Huizinga (1949) goes as far as asserting that play is not just an integral part of culture, like narratives and storytelling. It is more: play presupposes human society, and it is through play that animals and humans alike can learn. Bateson (1972) investigated how play-fighting monkeys were able to discern between their playtime activities like scratching and biting, and actual fighting, by communicating this meta-information about their type of activity to each other. This entails that play is something that can be entered wilfully by players. Importantly, the outcomes of playing do not have direct consequences for the world outside play (Salen & Zimmerman, 2004). When monkeys play, and, by extension, when humans play, they are aware of the boundaries of play and will act according to the agreed-upon rules of play. Through play, players are creating fictional play spaces in which they can experiment with their physical and mental skills. This way, players are able to attain and improve

NARRATIVE
IDENTITY

HOMO
NARRANS

NARRATIVE
RATIONALITY

NARRATIVE
COHERENCE

CHOOSE YOUR
OWN
ADVENTURE

INTERACTIVE
STORYTEL-
LING

HOMO
LUDENS

2 Here, I use 'reader' as a general term to indicate a person experiencing a story, whether in the form of a play, a film, or a different manifestation.

these skills in a safe environment. Despite Bateson asserting that play does not have direct consequences outside of play activities, the improvement of players' skills still transgresses those activities. Based on the notion that play underlies attainment of a lot of basic skills, especially in the early stages of life, there has been a massive increase in interest in creating playful learning experiences (Kirriemuir & McFarlane, 2004; Hirsh-Pasek, Golinkoff, Berk & Singer, 2008; Lillard et al., 2013).

In the following section, I go into more detail about using games to teach and learn. I choose the term *serious games* to describe the types of games I deal with in this dissertation. While being an indistinct term at first sight, serious games are not a *contradictio in terminis*, precisely because of Bateson's example of play-fighting monkeys (Bateson, 1955). Play affords learning and can hence be serious.

1.2 RESEARCH CONTEXT

The foundations of the domains of interactive storytelling and serious games can be found in the paradigms of *homo narrans* and *homo ludens*. The research I present in this dissertation builds on recent advancements in these two domains. Specifically, it focuses on the use of digital technology to improve both domains. In this section, I provide a short overview of these fields. In later chapters, I discuss interactive storytelling (Chapter 2) and serious games (Chapter 5) in more detail.

1.2.1 Interactive digital storytelling

The most basic example of making it possible for people to interact with a story is that of presenting a choice to readers at a certain point in the narrative through which they can decide how the story progresses. For example, a reader may choose to let the protagonist of a story take the road across high mountains or the road through murky swamps to reach her beloved prince. Depending on this choice, different events may befall the heroine. Extending this example, a story can be constructed that branches at several points, resulting in a larger number of story lines (Lethbridge & Mildorf, 2003). Such a *branching narrative* is typically seen in Choose Your Own Adventure novels such as the *Lone Wolf Series* (Dever & Chalk, 1984).³ In these novels, readers can influence how the plot proceeds by choosing from different options that are presented to them through text. One drawback of this method of using static decision points is that each storyline needs to be authored manually. Instead of previously having written one narrative, an author now has to write multiple coherent narratives. Because of this combinatorial explosion of story lines, this task becomes infeasible when authoring for a multitude of choices.

Linear stories as presented in most novels and films always feature a plot over which their authors have complete control. While opinions disagree on the precise definition of interactive storytelling (IS), the basis of all definitions resides in the notion that a story can be influenced by its reader (Murray, 1997). As this influence on the story goes beyond the typical affordances of a reader, I use

3 See <https://www.projectaon.org> for digitally augmented versions of these novels.

the term *interactant* as a description of a person interacting with a story. In her well-known work *Hamlet on the Holodeck*, Murray (1997) describes contemporary ideas about interactive fiction. Her vision of future interactive storytelling systems closely resembles the idea of *Star Trek's Holodeck*. She envisions the *Holodeck* as the ultimate form of interactive storytelling: it is a virtual reality environment which is indistinguishable from real life, including interactions with other (virtual) entities. Murray asserts that the feeling of *immersion* is important to experiencing an engrossing interactive story. This feeling is mostly constituted by providing interactants with sufficient *agency* so that they feel that their actions influence the story (Mateas & Stern, 2003). This introduces a challenge for the story's author: she has to design the system with which the interactant interacts in such a way that it is able to cope with the users' choices. In other words: how can a story be told while at the same time being interactive? This question captures the essence of the *narrative paradox* there is a trade-off between the plot of a narrative and the freedom of choice an interactant has (Aylett, 1999). The crux of the narrative paradox lies in the way a system balances the maintaining of a good plot and the provision of choices. The above example of a branching narrative in CYAO novels is only one possible approach to balance the possibility for making choices and having a structured plot. Such an approach only provides limited functionality to make a story interactive. To overcome this hurdle, the field of *interactive digital storytelling* (IDS) investigates methods that automate story generation and handle interactant actions (Arinbjarnar, Barber & Kudenko, 2009; Roberts & Isbell, 2008).

There are two main approaches in the field of IDS that try to tackle the narrative paradox, namely the *strong story* approach and the *strong autonomy* approach (Mateas & Stern, 2002). These approaches differ because of their focus: interactive storytelling systems (ISS) adopting the strong story approach aim to maintain a consistent storyline, while ISSs following the strong autonomy approach emphasise the autonomy of the characters in the story. Generally, a strong story ISS uses a top-down approach to steer a story's characters based on a narrative structure. This approach is most alike to the typical form of directing a play or a film: one person is in charge of the behaviour of all the others, instructing them about which actions to take. A possible consequence of this is that characters might exhibit inconsistent behaviour when they are directed to do so in the service of the plot. In contrast, a strong autonomy ISS uses a bottom-up approach by letting the characters interact and create the story through their interaction. While the strong story approach is rooted in practices such as theatre, the strong autonomy approach builds on those found in *improvisational (improv) theatre*. In improv theatre, actors receive instructions about an imaginary setting, for example, a location or a character description, and do their best at improvising an entertaining scene with this information (Johnstone, 2012). In strong autonomy ISSs, *virtual agents* take on the role of characters in a story. Virtual agents are autonomous entities in a virtual environment that observe what happens in their environment and act on this. The strong autonomy approach gives the agents more freedom to carry out actions than the strong story approach, but the former may result in a narrative without a coherent structure or overarching plot.

INTERACTANT

AGENCY

NARRATIVE
PARADOXINTERACTIVE
DIGITAL STO-
RYTELLINGSTRONG
STORY
STRONG
AUTONOMY
INTERACTIVE
STORYTEL-
LING
SYSTEMIMPROV
THEATREVIRTUAL
AGENT

In this dissertation, I investigate the field of interactive digital storytelling systems from two perspectives: that of the users (interactants) of such systems and that of the agents used in these systems. My work expands the research carried out with the *Virtual Storyteller* (Swartjes, 2010), a story generation system which has been adapted to enable user interaction (Alofs, Theune & Swartjes, 2015). I look at how users interact with the Virtual Storyteller and how agents can reason about a story's plot to overcome the narrative paradox. What I conclude is that both human and virtual agents can reflect on their interactions in and outside of the story and that we can use insights into the cause and effect of these reflections to improve the interactions with interactive storytelling systems.

1.2.2 *Serious games*

Over the years, a variety of terms has been coined to describe playful learning experiences, among others: simulations, simulation games, serious games, educational games, applied games, (digital) game-based learning, gamification, and edutainment (Breuer & Bente, 2010; Djaouti, Alvarez & Jessel, 2011). In this dissertation, I use the term *serious games*. I first provide an overview of various possible definitions of serious games before I propose the definition I use in the remainder of this dissertation. It turns out that the field of serious games is still struggling with a definition that is both broad and specific enough to contain its avails (Breuer & Bente, 2010; Djaouti, Alvarez & Jessel, 2011). This is also found in Abt's definition of serious games:

'Reduced to its formal essence, a game is an activity among two or more independent decision-makers seeking to achieve their objectives in some limiting context [in single-player games, the game or system is the opposing player]. A more conventional definition would say that a game is a context with rules among adversaries trying to win objectives. We are concerned with serious games in the sense that these games have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement.' (Abt, 1987)

Thus, in their broadest definition, serious games can be defined as all games that do not have entertainment as their prime purpose. These can be games that are intended to instil awareness in people, for example, about their influence on the environment. Serious games can be used for in-company trainings, such as decision-making role plays, or for practising fine motor skills, such as surgical accuracy. They may also be used for advertisement, scientific exploration, and simulating social processes. Therefore, this definition includes games used to educate or inform people (Djaouti, Alvarez, Jessel & Rampnoux, 2011), persuade them (Bogost, 2007), or assist in understanding scientific systems, such as chemical processes (Cooper et al., 2010). Additionally, this definition imposes no limitation on the implementation of a serious game. Serious games may, for example, incorporate state-of-start technology using virtual reality (Lohse, Hilderman, Cheung, Tatla & Van der Loos, 2014) or photo-realistic environments (Albert,

Hallowell, Kleiner, Chen & Golparvar-Fard, 2014). Yet they can also be analogue board games (Amaro et al., 2006) or educational role plays (Nestel & Tierney, 2007). With this broad range of application domains, taxonomies for serious games are equally wide-ranging. The taxonomy of Laamarti, Eid and El Saddik (2014) classifies serious games by their application area, activity, modality, interaction style, and environment (or interface). The Gameplay/Purpose/Scope (G/P/S) model of Djaouti, Alvarez and Jessel (2011) distinguishes between a serious game's gameplay, the rules that constitute how the game is played; its purpose, broadcasting a message, training cognitive or motor skills, and exchanging data; and its scope, the application domain.

Any definition of serious games presupposes an understanding of the definition of games themselves. Salen and Zimmerman (2004) give an explicit definition of games as follows: 'A *game* is a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome.' This definition can be complemented by that of Caillois (2001), who lists the following characteristics of games:

FUN The activity is chosen for its light-hearted character.

SEPARATE It is circumscribed in time and place.

UNCERTAIN The outcome of the activity is unforeseeable.

NON-PRODUCTIVE Participation does not accomplish anything useful.

FICTITIOUS It is accompanied by the awareness of a different reality.

For this dissertation, key words in these attempts at defining games are *artificial conflict* and *different reality*. A game is a game because it takes place by itself, in a world separated from the outside world. Salen & Zimmerman popularized the term *magic circle* to describe this separation, based on work by Huizinga (1949) and Caillois (2001). The play world or game world that players inhabit while playing is contained within this magic circle. It is a different reality: not a real one, but a fictitious, virtual one. This way, a game can offer players a safe environment to practice and explore their mental and physical skills.

MAGIC CIRCLE

Thus, the seemingly oxymoronic term of serious games starts to make sense. What is key to all games, serious or not, is that they convey information that can only be understood through the interactive nature of games. As van der Spek (2011, p. 4) puts it: 'All games, whether intended for leisure or "serious", revolve around learning.' The distinction between games for leisure and games with educational purposes lies in the use of the knowledge gained from the game. Contrary to entertainment games, knowledge gained in serious games is not only necessary to progress within the game's magic circle. This attainment of knowledge is also of importance outside the game's reality. In this respect, the goal of serious games lies not only in the experience while playing the game, but also in the experience when the game has finished and players have attained knowledge that is of importance in their real-world lives. Thus, in order to learn from serious games, players should be able to breach the magic circle and use their newly gained knowledge outside of the virtual world. This directly opposes Caillois' view that games should be

non-productive; serious games are only serious games when they actually are productive outside of their virtual worlds.

To narrow down the use of the term serious games in this dissertation, I define serious games as follows.

SERIOUS GAME A game through which players gain knowledge, skills, or awareness that is of importance outside of the game itself.

In this dissertation, I focus on serious games that have as their explicit learning goal educating people about social interaction, social awareness and related social skills. To this end, I was involved in the design, implementation and evaluation of several prototype serious games. These games relate to training certain cognitive and behavioural skills within the domain of law enforcement, specifically street interventions by the police.

Interactive systems such as those used for interactive digital storytelling and serious gaming always revolve around influencing what happens in the game or story world. In the context of this dissertation, the systems I study revolve around influencing the behaviour of virtual characters. Therefore, a part of this dissertation builds on research in artificial intelligence (AI) in order to automate the behaviour of these characters. Yannakakis and Togelius (2014) distinguish ten research areas of artificial and computational intelligence that were identified at a Dagstuhl seminar: non-player character behaviour learning, search and planning, player modelling, games as AI benchmarks, procedural content generation, computational narrative, believable agents, AI-assisted game design, general game AI, and AI in commercial games. My research in serious games for social skills training relates to both believable agents and computational narratives.

The main part of my research on serious games concerns ensuring people learn from these games. Therefore, I investigate how learning theories align with game design. Arnab et al. (2014) provide an extensive mapping of learning mechanics such as reflection and feedback to game mechanics such as role playing and assessment. Wu, Hsiao, Wu, Lin and Huang (2012) provide an overview of existing learning theories and how they relate to design practices for serious games, concluding that the latter often lack a proper underlying learning theory. The main learning theory on which I base my research is that of reflection with the help of feedback. The Serious Games Lemniscate Model (SGLM) schematically illustrates this concept, see Figure 1 (Koops & Hoevenaar, 2012). When someone interacts with a serious game, she enters a (play) cycle in which she observes what happens and processes this information to decide which action to take next. In order to learn, she needs to actively reflect on her actions and the game's responses by translating the in-game knowledge to real-world knowledge and then returning to the game world again (the learn cycle). Through this process of active reflection, this player takes a meta-perspective by breaching the magic circle's barrier. This illustrates how and why the meta matters for serious games. In this dissertation, I look at how this concept can be implemented in serious games for social skills training.

ARTIFICIAL
INTELLIGENCE

SERIOUS
GAMES
LEMNISCATE
MODEL

Figure 1.: The Serious Games Lemniscate Model (adapted from Koops and Hoevenaar (2012)).

1.3 GOING META: AGENTS REPUTANS

The fields of interactive storytelling and serious games have overlapping interests: the use of (interactive) narrative, the presence of intelligent agents, and the possibility of users reflecting on their experiences. My main research question boils down to this:

How can both users and virtual agents use meta-thinking to enrich interactions with interactive storytelling systems and serious games?

In these contexts, enriching interactions means improving players' interactions with these systems, for example, their engagement while creating a story and their learning gains when they play a serious game. In this dissertation, I explain why one of the approaches to answering this question is: by *going meta*. I do so by discussing the design of techniques that allow both virtual agents and users to go meta as well as by explaining how users experience their interactions from a meta-perspective.

GOING META

In their discussion of Huizinga's and Salen and Zimmerman's ideas about the magic circle, Frissen, de Lange, de Mul and Raessens (2015) do away with the notion that this circle is an impenetrable barrier. On the contrary, they emphasize that Huizinga never intended this term to mean that there is such a strict separation between the virtual and the real world. Instead, it is the case that players of games and readers of stories can be completely aware of inhabiting both worlds at the same time. This entails that user interaction is not limited to the virtual world: users may go meta and, for example, reason about the system's usability or the experience they are having. The Serious Games Lemniscate Model asserts that serious games should be designed with this exact thought in mind: players need to become aware of what they have been doing by reasoning on a meta-level about their experience in the game in order to learn from it. Thus, the game mechanics

of serious games need to be designed in such a way that they afford players to go meta and learn.

Similarly, players of interactive storytelling systems want to create and interact with stories that appeal to their fantasies and their desires. For example, some might want to explore a rich and varied story world, discovering hidden treasures, while others will be more into building their characters and having social interactions with others.⁴ This meta-level, which revolves around the players' personal experiences, requires interactive storytelling systems to take players' real-world desires into account, in addition to the their character's desires in the virtual story worlds.

Investigating how and why users go meta, and how to encourage them to go meta is essential to improving the experiences they have. Yet the story of this dissertation does not stop there. I assert that not only the users, but also the systems themselves need to go meta and deliberate about their situation. Like human agents, virtual agents can also go meta by thinking about ways in which a player's experience can be improved and adapting their system to match the user's needs.

Figure 2 shows a schematic overview of the interaction between players and agents, including their deliberative meta-cycles. The main interaction cycle in a virtual world (of an interactive storytelling systems or serious game) revolves around the player and the agent (the system) responding to each other through their respective sense-think-act cycles. That is, both types of entities sense the virtual environment through their available sensors, reason about the effects of the changes in the world they inhabit, and act towards goals or intentions. Such interactions take place in the virtual world, which is contained by the real world. These are the interactions which typically take place within the classic view of the magic circle. However, as explained above, both users and agents can meta-think about their interaction and take that into account when planning their actions. More than simply thinking about which action to take next, this process of meta-thinking lets both agent and user step outside of their role in the interaction. For example, a human player might think she doesn't like the story she is playing and try to take action to make it more enjoyable or simply quit the interaction. On a similar level, a virtual agent in a serious game may detect that the user has trouble with progressing to the next level, and adapt its behaviour to make it easier for her to continue.

To describe this paradigm of users and agents meta-thinking about their behaviour, inhabiting worlds both virtual and real concurrently, I coin the term *agens reputans*. Similar to *homo narrans* and *homo ludens*, describing man as storyteller and player, *agens reputans* denotes the concept of an entity (from Latin *agens*) being able to reflect and deliberate (from Latin *reputare*). Rather than being limited to humans (*homo*), this term explicitly emphasizes that any entity capable of acting (*agens*) on its own behalf can take a meta-point of view, and reflect and deliberate. In the following section, I explain how the research presented in this dissertation is supports this paradigm.

AGENS
REPUTANS

4 Cf. Bartle's work on player types (Bartle, 1996).

Figure 2.: A Venn diagram showing how players and agents relate to the virtual and real worlds. Both entities follow the same sense-think-act cycle and have the means to think on a meta-level about their interaction.

1.4 RESEARCH QUESTIONS

In this dissertation, I take an explorative approach to investigate the *agens reputans* from the two sides involved in human-media interaction: from that of the users of a system and from the side of the system itself, divided over the fields of interactive storytelling and serious games. This results in the following four research questions which form the core of this dissertation. The first two deal with meta-thinking in storytelling and the second two with meta-thinking in serious gaming.

RQ1 *What kind of meta-thinking do people do while co-creating stories?*

I investigate to what extent and about which factors players of interactive storytelling systems go meta. I address this question by analysing the communication of pairs of players who interacted with the *Interactive Storyteller*. This provides information about the design of the system, for example, whether its mechanics are clear and whether users are satisfied with their influence on the story.

RQ2 *How can virtual agents use meta-thinking to improve the generation of stories?*

I explain the design of techniques that enable virtual agents to deliberate about the story state to plan towards a good plot. These techniques let agents go meta and reason about which situations in the story may cause conflict, thereby enabling them to create more interesting stories. To exemplify this, I discuss how several such techniques were implemented in the *Virtual Storyteller*.

RQ3 *How can serious games stimulate meta-thinking to increase their effectiveness?*

Building on the notion that people do indeed meta-think about their interaction (RQ1), I designed techniques intended to stimulate reflection of players of serious games. These techniques let virtual agents show meta-information about the players' behaviour, which would otherwise not be available to them. This additional information is intended for players to go meta, letting them enter the reflection cycle of the Serious Game Lemniscate Model, and thus improving the learning effects of the serious game. I evaluate these designs in LOITER, a series of serious game prototypes revolving around gaining social awareness.

RQ4 *How can virtual agents use meta-thinking to adapt serious games to the user's competence?*

I describe how we designed techniques to provide adaptive difficulty in LOITER, based on the insights gained from answering RQ2 and RQ3. The game keeps track of the competence of players and makes this information accessible to the game's virtual agents. This way, the virtual agents change their behaviour to adapt the difficulty of playable scenarios, depending on this meta-information. To keep their behaviour coherent, the virtual agents provide a reason (within in the game world) why they have changed their behaviour. I discuss implementations of these techniques in a LOITER serious game, and analyse to what extent it influenced players' experiences and the game's learning effects.

1.5 OUTLINE OF THIS DISSERTATION

This dissertation is divided into four parts, of which Parts II and III tackle the research questions posed above.

PART I INTRODUCTION: SETTING THE STAGE

This is the part you, the reader, are currently reading. In this part, I introduce the context, research questions, and goals of this dissertation.

CHAPTER 1 THE NARRATIVE OF THIS DISSERTATION

In this chapter, I give a short overview of the fields of interactive storytelling and serious games. Four research questions, divided over the domains of interactive storytelling and serious games, treat the perspective of human and virtual agents going meta.

PART II META MATTERS IN INTERACTIVE STORYTELLING

This part illustrates how users and agents go meta when interacting with interactive storytelling systems.

CHAPTER 2 INTERACTIVE STORYTELLING

I start this part with a detailed overview of existing interactive storytelling systems and a discussion of the *Interactive Storyteller*, an interactive storytelling system which I used for further analysis.

CHAPTER 3 THE USER PERSPECTIVE ON INTERACTIVE STORYTELLING

This chapter addresses RQ1. I discuss how previous research has looked at interactions with interactive storytelling systems and I provide analyses of the way users meta-think while interacting with the *Interactive Storyteller*.

CHAPTER 4 AGENTS THAT PERCEIVE AND ASSUME TO STEER STORIES

This chapter addresses RQ2. I describe the workings of techniques for the *Virtual Storyteller* that enable virtual agents to meta-think to plan towards more dramatic plots than beforehand.

PART III META MATTERS IN SERIOUS GAMING

Similar to Part II, this part discusses how users and agents go meta, but this time, in serious games.

CHAPTER 5 SERIOUS GAMES

In this chapter, I provide an overview of research on serious games, detailing learning theories and existing systems.

CHAPTER 6 FEEDBACK IN SERIOUS GAMES

This chapter addresses RQ3. I scrutinize how feedback can be presented in serious games and explain my approach in designing techniques to stimulate meta-thinking, and hence reflecting, in our serious games.

CHAPTER 7 ADAPTIVE SERIOUS GAMES

This chapter addresses RQ4. I describe how I used techniques inspired by Chapter 4 to implement adaptive difficulty in our serious games by letting virtual agents meta-think.

PART IV CONCLUSION: REFLECTING ON THE RESEARCH

In this part, I go meta and reflect on my dissertation.

CHAPTER 8 LESSONS LEARNT

In this chapter, I recapitulate the answers to the four research questions and summarize which lessons I have learnt.

CHAPTER 9 FUTURE WORK

I close this dissertation with an outlook on future work, addressing challenges and sketching how these can be tackled.

PART II

META MATTERS IN INTERACTIVE STORYTELLING

In this part, I look at how the meta matters in the field of interactive storytelling, from the perspectives of both human and virtual agents.

In Chapter 2, I provide an overview of existing interactive storytelling systems and an explanation of the Virtual Storyteller, the system which we used for our studies in the following two chapters.

In Chapter 3, I answer Research Question 1:

What kind of meta-thinking do people do while co-creating stories?

I show that users of the Interactive Storyteller, an interactive version of the VST, actively go meta and reflect on their experiences. Designers of interactive storytelling systems can profit from analysing the users' reflection process, as it provides them with valuable insights into the functioning of their systems.

In Chapter 4, I answer Research Question 2:

How can virtual agents use meta-thinking to improve the generation of stories?

I show that the virtual agents in the Virtual Storyteller can improve the generation of stories by going meta. Using information that is only available to them from a meta-perspective, the agents can reason towards conflicts that might otherwise not come to pass. This reasoning may be used to improve the richness and variety of story generation.

We all make choices, but in the end, our choices make us.

Andrew Ryan
Bioshock

2K Boston and 2K Australia (2007)

THIS chapter provides background material on interactive storytelling for Chapters 3 and 4. It comprises two parts. In Section 2.1, I provide an overview of interactive digital storytelling systems. In Section 2.2, I discuss the system I used and extended for this part of this dissertation: the (Interactive) Virtual Storteller.¹

2.1 INTERACTIVE STORYTELLING SYSTEMS

In Chapter 1, I discussed the origins of the field of interactive digital storytelling. Below, I provide an overview of a selection of well-known interactive storytelling systems (ISSs) by discussing their unique techniques and by comparing their merits and drawbacks. The systems developed in this field can be distinguished by their approach of steering the story: either through a centralized, top-down director agent (the *strong story* approach), or through multiple actor agents collaborating to create a story (the *strong autonomy* approach) (Linssen, 2012). First, I discuss the ISSs using the strong story approach in Section 2.1.1 and ISSs using the strong autonomy approach in Section 2.1.2. In Section 2.1.3, I discuss two systems which use a combination of the two approaches.

2.1.1 The strong story approach

As described above, strong story ISSs emphasize sustaining a coherent plot over the consistency of the behaviour of individual characters. These ISSs rely on planning and re-planning of such story plots. I describe two such systems that use techniques to overcome the narrative paradox.

2.1.1.1 Mimesis

Designed as a module for use in existing game engines, the *Mimesis* architecture is dedicated to the creation of narratives in games (Young et al., 2004). It is a *drama manager* which is responsible for the generation of plans for a game and its agents. It safeguards the coherence of generated stories when situations arise

MIMESIS

DRAMA
MANAGER

¹ This chapter is based on Linssen (2012) and Swartjes (2010).

that deviate from the constructed plan. In Mimesis, narratives are first constructed as plans, which are triggered by plan requests by the game itself in the form of a certain goal or state the story needs to reach, see Figure 3. Plan-construction is a three-stage procedure that consists of story world planning, discourse planning and execution management. The story world planner composes a plan that fulfils the given goal through use of available actions, based on a certain world state. This plan is fed forward to the discourse planner, which combines the story world plan and its associated narrative elements, for example, background music. Finally, the execution manager constructs and monitors a directed acyclic graph of the complete plan. Each action that is to be carried out in the game world represents one node in the graph. When such an action is executed, the corresponding node is removed from the graph and the narrative advances according to plan.

Figure 3.: The Mimesis architecture, using Unreal Tournament 2003 as a sample game engine (adapted from Young et al. (2004)).

NARRATIVE MEDIATION

In ISSS, it is rule rather than exception that a player does not behave according to any pre-constructed plan. In this case, Mimesis uses *narrative mediation* to stay one step ahead of the player by detecting actions that may result in possible conflict with one or more steps of the plan (Riedl, Saretto & Young, 2003). Riedl et al. developed two approaches to narrative mediation that either try to accommodate the plan or intervene, circumventing minor and major conflicts, respectively. When accommodating, Mimesis allows the player's action which is an exception to the current plan to be carried out, but revises the remainder of the narrative plan to account for conflicts of the action. Intervening on the player's exceptional action substitutes its expected outcome by one or multiple effects that do not threaten the constructed plan. Riedl et al. give the example of a player trying to purchase a beverage from a vending machine for which she inserts a coin which has a different intended use in the game. Therefore, the game intervenes and lets the vending machine produce an error so that the coin is not accepted and given back to the player in order to be used for its intended goal. It is important that the effects of interventions do not deviate too much from the player's expectations. For instance, if the same logical action (inserting the coin) fails to produce its usually expected effects (receiving a drink) again and again, this can lead to a decrease in the player's agency.

Narrative mediation has two limitations to its use. One of these lies in the recognition of exceptional actions, because they may also arise as steps in a longer plan of the player and not merely as single constituents of one. In that case, exception recognition needs to expand its search space in plans, which may result in feasibility problems. Second, intervention has only been designed to replace the effects of one action. In theory, narrative mediation seems a useful technique for coping with the narrative paradox. However, its reliance on detecting deviations from plans and its need for authoring for these exceptions imply feasibility problems that have to be overcome before it can be implemented successfully.

2.1.1.2 U-Director

Adopting a detective-style setting, *Crystal Island* is an educational 3D game in which the player tries to determine the cause of illness in his team of scientists (Mott & Lester, 2006). She does so by inspecting objects, witnessing events and through dialogues with virtual characters. To construct a storyline with a logical succession of events, an architecture called the *U-Director* is used (see Figure 4), which is designed to maximise the narrative utility of the story. Mott et al. describe this as the game aiming to give the player sufficient hints about where she should go next or what actions she should perform. These hints should not be overly explicit to prevent the player’s feeling of agency dropping or losing the overall challenge of the game.

CRYSTAL
ISLAND

U-DIRECTOR

Figure 4.: The U-Director architecture (adapted from Mott and Lester (2006)).

U-Director’s reasoning is based on *dynamic decision networks*, which are Bayesian networks extended with decision processes that incorporate temporal aspects. At each moment in time (a time slice), the U-Director keeps track of the plot progress and the narrative flow (the narrative objectives); the plot focus and the physical

DYNAMIC
DECISION
NETWORK

state of the story world; and a model of the user's goals, beliefs and expectations (the user state). Effects of director actions and how they influence the narrative are calculated for the next time slice as well as the effects of possible user actions for the time slice thereafter. The dynamic decision networks evaluate the available actions to select the director action that enables maximal narrative utility after which the steps are repeated for the successive time slices.

Maximizing the narrative utility is analogous to making the storyline conform to a desired story arc. Unlike Mimesis, U-Director's approach does not directly intervene on a player's actions, but tries to nudge her toward a desired state from which the story can progress.

2.1.2 The strong autonomy approach

Strong autonomy ISSs place the responsibility for handling the narrative paradox into the hands of a story's characters. These ISSs rely on characters who are modelled in more detail than those in strong story ISSs, with their interaction with a player determining how a story progresses. In this section, I discuss only one strong autonomy approach; I expand on a second one, the Virtual Storyteller, in Section 2.2. Other approaches that may be considered strong autonomy approaches often are hybrid approaches as they use a centralized drama manager to structure the overarching story (see Section 2.1.3).

2.1.2.1 EmoEmma

EMOEMMA *EmoEmma* is an interactive storytelling system focusing on immersion. It was designed to be a virtual reality demonstrator for multi-modal interaction in a nineteenth century setting in which the player interacts with Emma Bovary, see Figure 5 (Cavazza, Lugin, Pizzi & Charles, 2007). It takes its subject matter from Gustave Flaubert's novel *Madame Bovary* (Flaubert, 2004), in which Emma is a lady of blue blood who is fed up with her lifestyle and decides to run away. EmoEmma's scenario is mainly influenced by a few chapters in which Emma decides to have an affair with Rodolphe, another aristocrat. During the game, the player acts as Rodolphe and can interact with a virtual character representing Emma who does not have a concrete goal, but a set of feelings that she wishes to reach (Pizzi & Cavazza, 2007). This is in line with Flaubert's thoughts about his novel that the narrative should focus on feelings instead of on actions.

MADAME
BOVARY

EmoEmma is played in a virtual reality environment with life-size in-game characters. Physical interaction is limited to recognition of the distance of the player to the characters and whether or not the player is facing a character. A player can influence the story by either agreeing or disagreeing with Emma which leads to different outcomes. For example, when Emma says that she 'would leave everything behind for you [the player],' the player can turn his back to her which is recognised by the system and to which Emma responds by expressing disappointment.

This ISS distinguishes itself from the previous systems in that it lets virtual characters infer how to feel through a heuristic search planner. Pizzi and Cavazza

Figure 5.: The reasoning behaviour of agents in EmoEmma (adapted from Cavazza, Lugrin, Pizzi and Charles (2007)).

(2007) give the example of when a character needs to feel bored, which can be recognised by determining the ‘distance’ of a character to his or her intended emotional state. For example, if there has been a lack of progression for some time, the character will recognize this as boredom and will start to perform an action to overcome it. This type of inference is performed by the interpretation operators, which update the character’s feelings. The characters have two other operators at their disposal: character interaction operators, which modify another character’s mental state; and physical operators, which are used for physical actions. EmoEmma lacks a drama manager, as each character is motivated purely by his or her own feelings.

Lugrin, Cavazza, Pizzi, Vogt and André (2010) evaluated player involvement in this multi-modal experience in a between-subjects experiment. In one condition, the participant assumed the role of Rodolphe and was able to interact with Emma directly. In the other condition, the player was a ‘ghost actor’ who could influence the characters by manipulating the physical world and by influencing their thoughts through speech. Results from the experiment showed that both conditions resulted in equal story lengths as well as equal amounts of user actions. This is an indication that both approaches to user interaction may be equally suited for further use in ISSs.

In essence, EmoEmma adopts an approach that is much more concerned with a ‘*narrative* description of characters’ psychology rather than *cognitive* models’ (Pizzi & Cavazza, 2007, p. 116). That is, the behaviour of the agents conforms to narrative norms and expectations, but does not necessarily entail natural, real-life behaviour. This makes it suitable for story generation, yet less so for more realistic scenarios which emphasize coherent behaviour of the characters from a cognitive point of view.

2.1.3 Hybrid approaches

Some systems use a combination of strong story and strong autonomy techniques, letting drama managers share responsibility for a good plot together with autonomous characters.

2.1.3.1 FearNot!

FEARNOT! Fun with Empathic Agents to Reach Novel Outcomes in Teaching (*FearNot!*) is a demonstrator devised for children to evoke awareness of bullying practices and to teach them how to cope with such behaviour (Aylett, Dias & Paiva, 2006). Users assist the game's protagonist through different scenarios by giving advice to him after conflict with a bully. There is no direct intervention by the player as the protagonist and the bully behave according to their emotions during these episodic conflicts. In each of these scenarios, the bully assumes a dominant stance and tries to overpower the protagonist, for example, by inducing fear. If users give good advice to the protagonist, he may overcome his fear of the bully and stand up to him in successive scenarios.

FATIMA FearNot!'s agent architecture *FATiMA* (FearNot! Affective Mind Architecture) enables the characters to be subject to emotions by implementing a subset of the OCC model, a framework for emotions (Ortony, Clore & Collins, 1990). *FATiMA* specifically features hope and fear, both of which are emotions with a temporal aspect which can influence the creation of plans. When a character is driven by hope, he may become bolder and more optimistic, whereas fear may let characters refrain from carrying out certain plans. To generate emotions, *FATiMA* relies on the appraisal of and coping with events, see Figure 6. *FearNot!* also incorporates decay of the intensity of emotions over time because of the dynamic aspect of emotions.

Figure 6.: The FearNot! Affective Mind Architecture (adapted from Aylett, Dias and Paiva (2006)).

An important property of FAtiMA is that its appraisal mechanism is comprised of a reactive layer and a deliberative layer, accounting for instantaneous emotional responses like anger, and prospective ones such as hope and fear. The deliberative layer creates and changes plans based on the appraisals of every event it witnesses and checks what the effects are on the character's goals. The same distinction between reactive and deliberative behaviour is found in the character's coping mechanism, in which reactive actions are triggered when certain conditions are met. The goal of deliberative reasoning is to generate the strongest emotion, as Aylett et al. assume this to be the most urgent one. FAtiMA generates a plan that will fulfil this goal, after which the emotions of hope and fear are updated, corresponding to that plan. These emotions denote how likely a character thinks his or her plan will succeed based on the possible effects on his or her current state. Furthermore, deliberative reasoning has the possibility to use one of three emotion-focused coping strategies: acceptance, acknowledging a plan failure; mental disengagement, lowering one's expectations of a plan; and denial, ignoring possible threats to the plan.

Aylett et al. (2008) propose two approaches to overcome the narrative paradox. The first of this is inspired by *plot* or *way points* as used by Weyhrauch (1997). The *Story Facilitator* structures the overarching narrative into different episodes based on the actions taken by the characters in previous episodes. This resembles a *strong story* approach, as a central entity decides the confines within which character agents have their say. This control does not influence the characters in each of the episodes directly, but provides a context in the following episode that, for example, features characters with whom the player had interacted before, and in which location the episode takes place.

PLOT POINTS
STORY
FACILITATOR

Out-of-character thinking (OOC) is the second approach to dealing with the narrative paradox in FearNot! uses a strong autonomy approach. The characters are designed in such a way they are aware of the dramatic impact of their actions through a process called *double appraisal*. This process determines all the possible emotional consequences of the effects of available actions and ranks these according to their emotional impact. Then, the action with most emotional and therefore dramatic impact on all characters is selected to be carried out. Here, we see the possibilities of characters that think beyond their role in the story. In Section 2.2, I expand on the possibilities of OOC reasoning in an ISS called the Virtual Storyteller.

OUT-OF-
CHARACTER
THINKING

DOUBLE
APPRAISAL

2.1.3.2 Façade

Perhaps one of the most well-known interactive storytelling systems is *Façade* (Mateas & Stern, 2003), a 3D 'interactive drama' which places the player in the role of a person visiting a couple (Grace and Trip) whom she has not spoken in a while.² This is all the background information that is given to the player at the start of the game. Mateas and Stern's goal was to create a narrative that evolves over time based on the player's interactions with the characters and the story world. The player can interact with the couple by talking to them through text

FAÇADE

² Façade is available at <http://www.interactivestory.net>.

input as well as by performing physical actions through mouse clicks. Through interacting with the couple, she can discover and influence what is going on between them. For example, the player may ask the couple how they are doing after which they respond a little bit hesitantly that everything is okay. If the player decides to continue questioning them it may become clear that they in fact have some marital problems. Then, she can choose to try to comfort them or to take sides in the ensuing discussion. There are no static points at which the player has to choose distinct options as everything happens continuously and the player is able to interrupt the couple when they are saying something.

Central to *Façade*'s design is a drama manager that manages the narrative progression, see Figure 7. It does so by trying to mould the narrative so that it corresponds to a desired story arc such as the one of *Freytag's Pyramid*, see Figure 8. Freytag divided dramas into five parts: exposition, rising action, climax, falling action, and dénouement, as is explained in the discussion of Freytag (1863) by Lethbridge and Mildorf (2003). Exposition implies an introductory state, wherein the reader is familiarised with the story's subjects. Rising action describes the events that occur and actions that are undertaken previously to the story's climax, while building the tension. The climax features the highest peak in dramatic tension. Thereafter, there is falling action, namely the decrease of tension by letting the main conflict unravel. Lastly, dénouement addresses the resolution of the conflict and the final release of tension.

Figure 7.: The *Façade* architecture (adapted from Mateas and Stern (2003)).

Façade uses *story beats* to manage the narrative.: instantiations of story situations, for example, one in which Grace and Trip both ask the player what he wants to drink. These beats have preconditions and effects that serve the purpose

Figure 8.: Freytag's Pyramid, a model of the dramatic tension in narratives.

of a partial ordering of a collection of beats. By sequencing beats based on their correspondence to the story arc, a narrative of the desired form emerges. Beats become either available or unavailable based on a player's actions, which are first translated into *discourse acts*.³ These are abstractions of the player's actions, for example, 'praising' or 'expressing anger'. This happens through natural-language processing, which categorizes the meaning of the player's action, after which a potential reaction is chosen based on the beat the story is in. In this manner, the player is able to influence the story's flow and outcome. Each beat enables certain behaviour on the agents' behalves, for when the player seems to be expressing anger, Grace and Trip need to respond in a natural way. For example, they can react somewhat laconically and ask if the player really meant to react this way, or they can point her to the door if they feel they are tired of her behaviour.

Mateas and Stern note that although their framework enables an emergent narrative to be constructed, a large amount of effort was put into the creation of the characters' behaviours (Mateas & Stern, 2005a). Responses to situations and utterances still needed to be authored, which called for envisioning possible situations. As Mateas and Stern (2005a) explain, this does indeed limit the interactivity of *Façade*, as only pre-constructed behaviours, specific to pre-envisioned situations, can be enacted by the characters.

2.1.3.3 IN-TALE

The Interactive Narrative Tacit Adaptive Leader Experience (IN-TALE) is a virtual environment dedicated to negotiation training for the military. It uses a similar approach to that of *Façade*: it combines the limited autonomous behaviour of *Façade*'s character with the incorporation of a drama manager alike to that of *Mimesis* (Riedl & Stern, 2006a). IN-TALE restricts the agents' behaviour to non-scenario-specific activities (*local autonomous behaviour*, LAB). Complementing this, IN-TALE's drama manager concerns itself with scenario-specific actions (*narrative directive behaviour*, NDB). More concretely, an example of a LAB is an 'opening the store' activity, which consists of all the subtasks necessary to achieve its effect. NDBs describe more complex behaviour that is tightly related to a specific scenario, for example, for a scenario involving a terrorist, 'planting a bomb'. These behaviours are used by a drama manager in the same way that *Mimesis* used

DISCOURSE
ACT

IN-TALE

LOCAL
AUTONOM-
OUS
BEHAVIOUR
NARRATIVE
DIRECTIVE
BEHAVIOUR

³ In this context, 'discourse' means the dialogue that is performed, not the way of representation.

its story world plans. The plans are evaluated before execution and checked for possible inconsistencies between the simulation state and the expected narrative. For each inconsistency, an alternative plan is created. However, this only leads to a shift in problems as pre-constructing all possible plans would still lead to a combinatorial explosion of possibilities, hence the need for a dynamically constructed narrative.

The conflict between character actions and planned goals is further elaborated on by Riedl and Stern (2006b), who describe agents that *fail believably*. To do so, the distinction between prescriptive and proscriptive directives is made, which instruct characters which goals to pursue and which story world states to refrain from reaching, respectively. When characters are given a prescriptive directive to achieve something, they can either merge their current goal with the given goal; complete their current goal and continue with reaching the given goal; or they can abort or suspend their current behaviour and focus on the given goal. In case of the latter situation, an agent uses a transition behaviour to make this transition and still remain believable.

Proscriptive directives entail analysis of the constructed plans in order to check whether a conflicting situation arises as well as failing believably when a conflicting, planned action is inevitable. A *restoration mode* serves as a safety net for the latter as characters are given the ability to explain why they refrain from continuing to perform conflicting actions. This differs from the above transition behaviour in that the character already is in conflict with a directive.

2.2 THE INTERACTIVE VIRTUAL STORYTELLER

The research I discuss in Chapters 3 and 4 builds on previous work done on the Virtual Storyteller and its derivative, the Interactive Storyteller. Here, I discuss the workings of these (interactive) storytelling systems.

2.2.1 The Virtual Storyteller

In 2001, Faas (2002) made the first attempts at designing and implementing *The Virtual Storyteller* (vst), a system meant to generate stories.⁴ Rather than generating stories through explicit representations of story structures, the vst adopted a character-based view. Inspired by theatre, the vst uses agents that take on different roles in the system, see Figure 9. Crucial to the creation of a story are the *Director* and *Character Agents*. The Director Agent issues story roles to the different Character Agents and makes sure that they have goals to attain in the story world. This way, it tries to guide the story by narrowing the possibilities of the Character Agents. These, on the other hand, attempt to plan and act out actions towards attaining the goals that have been set out for them. When a story has been acted out, the *Narrator Agent* translates the actions that have taken place into natural language text. Finally, the *Presenter Agent* is able to narrate this story using an embodied agent with speech capabilities.

⁴ See <http://www.virtualstoryteller.info> for a complete overview of all the work relating to the vst.

Figure 9.: A schematic overview of the agent-based architecture of the vST.

In Transcript 1, I show a story generated by Faas' implementation, reproduced from (Faas, 2002, p. 50). Despite being a relatively simple story, this example shows the processes underlying its generation in the vST. What is apparent at first sight is that the story is presented in natural language, courtesy of the Narrator Agent.⁵ At the beginning of the story, the current state of affairs is explained. This is followed by a motivation (the dwarf being hungry), which leads Plop to take up the goal of eating something to overcome his hunger. He does this by planning actions to go to the house as he knows it contains an apple that he can eat to satiate his hunger.

Once upon a time there was a dwarf. He was called Plop. Plop was in the forest. A house was in the forest. A apple was in the house. Plop was hungry. Plop walked to the house. He entered the house. He picked up the apple. He ate the apple. Plop lived happily ever after.

Transcript 1.: The story of a dwarf called Plop.

The crux of the vST is that, through the Character agents' interactions with each other and the story world, a story emerges. The vST originally started out as a hybrid approach to character-based story generation, with the Director agent guiding the Character agents in the stories. In the years following the creation of the vST, several adaptations have been made, most notably by Swartjes (2010). He improved the way the vST represents causality between actions and events in the story world, which is termed the *fabula* of a story (Bal, 2009). Swartjes revised the *General Transition Network*, a fabula model by Trabasso, Stein, Rodkin, Park Munger and Baughn (1992) used to describe subjective interpretations of characters in stories, to allow for objectively capturing events in stories. Figure 10 shows the fabula model designed by Swartjes and Theune (2006).

FABULA
GENERAL
TRANSITION
NETWORK

⁵ Slabbers (2006) and Hoek (2014) have since looked at improving the Narrator agent, enabling more proper use of grammar as well as the generation of more complex sentences.

Figure 10.: The vst's fabula model, showing the possible causal relations in the story world (adapted from Swartjes (2010)).

In total, there are seven types of elements in this fabula model, with four possible causal relations between them. I briefly discuss them here and refer to Swartjes (2010, pp. 123–124) for a more in-depth discussion of his model.

SETTING ELEMENT (SE) An initial state of some part of the story world before starting the system.

GOAL (G) A character's desire to attain, maintain or avoid some state of affairs.

ACTION (A) Any change a character can affect on the story world.

OUTCOME (O) Whether or not a character attained a goal.

EVENT (E) Any change in the world not directly resulting from a character's action.

PERCEPTION (P) The act of a character perceiving an action or event happening.

INTERNAL ELEMENT (IE) A character's conscious and subconscious thought process, such as emotions and beliefs.

As can be seen in Figure 10, these elements can be causally connected by either physical causality (ϕ) resulting from an event or action causing something, a character's motivation (m) through adopted goals, physiological causality (ψ) causally linking perceptions and beliefs, and enablement (e) which simply satisfies the preconditions of another element.

Character agents try to attain their goals, but it may be the case that they are unable to reach them, because they do not have an Action available to do so. The Director agent might have chosen a different story setting to prevent this from happening or it might try to change the story so that the problem is circumvented,

as we saw in the strong story approaches described earlier in this chapter. However, Swartjes, Kruizinga and Theune (2008) argue that the Character agents can solve these problems by giving them more responsibilities in creating stories. To do so, they looked at the way improvisational theatre (*improv* in short) tackles such situations. When improvising, actors in such plays are aware of their being in a story, which they can use to their advantage by planning actions in such a way that a coherent and meaningful story emerges from their interactions. This notion of meta-thinking in improv is called as out-of-character reasoning (ooc), equivalent to that in FearNot! (Aylett et al., 2006). Translating this idea to the vst, the Character agents can think on two different levels: in-character (ic) and out-of-character. Through this addition of the ooc level of thinking, Character agents can deliberate on which actions are of best interest to the story as a whole. For example, if the story seems to halt because the characters cannot progress towards their goals, they can use ooc reasoning to solve this.

Late commitment is a technique of using ooc reasoning implemented by Swartjes et al. (2008). It heavily relies on the idea in improv theatre that a story world is not fixed when a play is being acted out. Instead, the actors themselves build this world by offering new information, for example new locations that appear or new, unexpected events that take place (Johnstone, 2012). Moreover, improvisation wholly relies on stories being created this way, as late commitment can have dramatic effects on the actors and viewers of the play alike. Through late commitment, Character agents can decide to frame the story the way they would like to see it. For example, a pirate character wants to attain a goal of obtaining a pile of gold. Two *framing operators* enable him to attain it as they satisfy the goal's preconditions, see Figure 11. First, the Character agent asserts that there is gold in a mine, using the *GoldInMine* framing operator. However, no mine does yet exist in the storyworld. Using a second framing operator, *MineAtIsland*, the Character agent asserts that there is a mine on the island to fulfil this requirement and complete his plan of attaining the gold.

Ultimately, the vst was able to generate richer and more entertaining stories than its first incarnation, based on Swartjes' work. In Transcript 2, I show one such story from the *PLUNDER*⁶ domain which features the use of framing operators as well as more complex characters interactions and more advanced narration (Swartjes, 2010, p. 179). This story starts out with two characters with initially separated goals which they try to solve by themselves. Yet as soon as Billy sees that Anne drinks some rum, he also becomes thirsty for some rum and uses a framing operator (indicated by sentences in typewriter font) to incite a conflict between the two of them. The vst provides a way to generate stories with a strong autonomy approach, as the characters themselves are fully autonomous in choosing their actions. Yet, with the addition of framing operators, this system is also able to generate stories that have some sense of direction towards a good plot.

LATE COM-
MITMENT

PLUNDER

⁶ *PLUNDER* is an acronym of Pirates Looming in Unscripted Narrative: towards Dramatic Emergent Roleplay.

Figure 11.: Example schematic fabula. The pirate is able to attain his goal because he uses out-of-character reasoning to employ framing operators (adapted from Swartjes (2010)).

The water supply contains water.
Billy Bones is the captain of the ship.
 Suddenly, the ship's water supply is exhausted!
 Anne Bonney becomes thirsty for some rum.
 Billy Bones sets sail towards Treasure Island.
 Anne Bonney takes the bottle of rum from the deck.
 Billy Bones walks back and forth, replenishing the water supply with
 water from a pond of fresh water on the island.
 Anne Bonney drinks rum from the bottle of rum.
 Billy Bones becomes thirsty for some rum.
Billy Bones hates drunkards.
 'Prepare to meet your seaman's grave, Anne Bonney!' says Billy Bones.
Anne Bonney carries a rapier.
 'No way, Anne Bonney!' says Billy Bones.
 Anne Bonney draws a rapier.
Billy Bones carries a rapier.
 Anne Bonney stabs Billy Bones with a rapier.

Transcript 2.: The use of framing operators in the PLUNDER domain.
 Framing operators are shown in *italics*.

2.2.2 The Interactive Storyteller

In Section 2.2.1 on interactive storytelling systems, I outlined several such systems that use their own specific methods to handle user interaction with the story. While the Virtual Storyteller was not initially planned to lend itself for user interaction during story generation, the developments by Swartjes (2010) opened new possibilities. As I explained above, actors in improv theatre can adopt character roles to act out a story. Analogously, it should be possible for human users to adopt character roles in the VST, taking over the responsibilities of the Character agents. The main challenge in this was a technical one, as users need an interface through which to control their characters and interact with the story world. Based on the VST, Alofs et al. (2015) created a table-top system called the *Interactive Storyteller* (IST) which lets users act out a variation on the story of *Little Red Riding Hood* (J. Grimm & Grimm, 2008) in a story domain of the VST specifically manufactured for this purpose (RED). The target group of users for the IST was children aged 8 to 12 years. The main motivation behind creating this system was to enable co-creation of stories through a symbiosis of these users and virtual agents. All interactions with the IST took place through the Surface Touch, a multi-touch table, see Figure 12.

Figure 12.: The Surface Touch table, on which the Interactive Storyteller could be played.

While interacting with the IST in the RED domain, users can adopt any of three roles in the story: that of Red herself, of her Grandma, and of the Wolf. Assigning these roles can be done by selecting a character's portrait in the interface and choosing whether that character should be controlled by a user or an AI. Interaction takes place in a turn-based fashion, with each character in the story being allowed to take one action in their turn. When users control a character, their character's available actions are presented to them through the IST's interface as in Figure 13. Different actions are available to characters, depending on their occupied location, previous actions, other characters present at their location, and held items. For example, Red can walk from her house to Grandma's house

and hand her a delicious cake. Users can move characters between the adjoining locations, indicated by blue circles, by dragging their portraits from one location to another.⁷ If the Wolf chases Red and apprehends her, he can grab the cake from her hands to eat it himself. When this happens, Red and Grandma can conspire to bake a poisoned cake to get back at the Wolf, who will appease his hunger with the cake and die a tragic death.

Figure 13.: The panel through which users can decide which action to take. The actions are readable from both sides of the table.

When characters are not controlled by users, they will behave autonomously, creating their own plans based on goals they adopt. For example, the Wolf will always become hungry after a set amount of turns, leading him to adopt the goal of eating something to satiate himself. Therefore, he will determine whether there is something edible in the story world, like Red’s delicious cake. If he detects such a delicacy, he will plan to journey to the food’s location, retrieve it, and consume it. As alluded to above, Red’s initial plan is to get the cake to Grandmother. If she succeeds, Grandmother will be happy with the cake. If Red fails to do so, because the Wolf snatched her cake, she will converse with Grandma to form the plan of killing the Wolf with a poisoned cake. As the Wolf will become hungry again, he will surely eat the poisoned cake, leading to the completion of Red’s goal of taking revenge.

The main goal of the IST was to stimulate social interaction between multiple users while co-creating a story with the system. Interactions with the system were set up to include two children who played the roles of Red and Grandma, with a Character agent taking on the role of the big bad Wolf. Figure 14 shows the multi-user interface of the Interactive Storyteller that was used. To support the children’s exploration of the story world, several locations and actions are available to be visited and carried out. For example, characters can play on the beach or look at birds at a clearing in the forest. The interface narrates the character’s actions through the Narrator agent, which uses a text-to-speech module. Additionally, all actions are listed in natural text in story scrolls, which can be read, enlarged and scrolled through.

A series of interactions with the IST by four pairs of children yielded a variety of created stories (Alofs et al., 2015). Each of the pairs played two sessions, one with the touch-only version of the system, and one with the characters being

⁷ To increase the similarity to board games, Alofs et al. created another form of interacting with the characters, namely through tangible playing pieces instead of the portraits in the interface. These playing pieces can be recognised by the multi-touch table and can therefore be dragged from location to location in the same manner the children would drag the portraits.

Figure 14.: The graphical user interface of the Interactive Storyteller.

represented through tangibles. Play sessions typically lasted about 12 minutes, during which the interactions were recorded through video and the created stories. One such story, co-created by two children, is shown in Transcript 3.⁸

This log of actions does not show a complete picture of the interaction. The reasons for the children taking some actions in the story above only became clear when analysing the video footage. Grandma's seemingly random outburst in tears was explained by one child by saying that Red was leaving Grandma's location. Furthermore, this analysis showed that interacting with the IST incited various social behaviours between the children. For example, the children made lots of shared eye contact and laughed often. They also discussed how they could help each other's characters in the story and tried to plan their actions.

The evaluation of the IST by Alofs et al. (2015) raised several questions. Firstly, the coherence of the children's co-created stories is not always apparent at first sight. How do we determine whether the children were goal-oriented or were just playing around? Secondly, Alofs et al. analysed the children's social interaction by determining which kinds of behaviour they showed. Yet on which topics were the children discussing their experience? In Chapter 3, I scrutinize the children's interactions with the system in more detail by cross-examining the story logs and the video footage. I argue that, to more fully understand such user-system interactions, we need to look at them from a meta point of view.

⁸ This story has been translated from Dutch to English, see Transcript 19 in Appendix A for the original version.

Once upon a time, there was a little girl with a red cap. She wanted to bring a birthday cake to her grandmother..

Wolf is mean.

Red bursts out in laughter.

Grandma bakes the cheese cake.

Wolf sneaks to Red's house.

'Hello, Wolf,' says Red.

With a little bit of cyanide, Grandma poisons the cheese cake.

'Oh, hey, Red!' says Wolf.

Red does the dishes.

Grandma shuffles to the beach.

'Give me the birthday cake,' says Wolf, and forcefully takes it away from Red.

'Give that back!' says Red, and takes the birthday cake back from Wolf.

Grandma shuffles to Red's house.

'Give that back!' says Wolf, and takes the birthday cake back from Red.

'Oh Grandma,' says Red, 'Wolf stole the birthday cake from me!'

'Oh that is not nice, I feel very sorry for you,' says Grandma to Red.

Wolf eats the birthday cake.

'Here you go, Wolf,' says Grandma, giving the cheese cake to Wolf.

Red bursts out laughing.

Grandma bursts out in tears.

Wolf becomes very hungry.

'What should I do, Grandma?', asks Red.

'Well, Red,' says Grandma, 'I have a plan to poison Wolf. You just wait and see!'

Wolf eats the cheese cake.

Transcript 3.: A story created by two children with the Interactive Storyteller.

THE USER PERSPECTIVE ON INTERACTIVE STORYTELLING

3

Nobody gets a good view of a system from the inside, no matter who they are. The view's distorted.

Siri Keeton
Blindsight
Watts (2006)

3

IN this chapter, I discuss how the meta matters to users of the Interactive Storyteller (IST).¹ This chapter deals with Research Question 1 of this dissertation:

RQ1 *What kind of meta-thinking do people do while co-creating stories?*

To answer this research question, I investigate interactions of pairs of children with the Interactive Storyteller by looking at the coherency of their actions and the communication between these children during play. As described in the previous chapter (see Section 2.2.2), the Interactive Storyteller system lets users create stories in collaboration with virtual agents that respond to the users' actions. Intended for children around the age of 10, the IST uses the RED domain, which is a variation on the story about Red visiting Grandma and encountering the Wolf (J. Grimm & Grimm, 2008). A previous analysis of children interacting with the IST by Alofs et al. (2015) showed that children were able to use the system to co-create stories. My approach in answering the above research questions comprises two steps. Firstly, I discuss how we investigated interactions of pairs of children with the Interactive Storyteller by looking at the coherency of their actions. Secondly, following on this study, we looked at the communication between these children. This turned out to be paramount to understanding both their created stories and their interactions with each other and the IST. More specifically, the children actively conversed about their interaction from a meta-perspective, discussing their behaviour in the storyworld, the system itself and their thoughts about their created stories. In this chapter, I provide a method to scrutinize such interactions by classifying to what extent the children go meta when interacting with the IST.

In Section 3.1, I discuss previous work on the perspective of users on interactive storytelling systems. The remainder of this chapter is based on the interaction corpus collected by Alofs et al. (2015), which I discuss in Section 3.2. In Section 3.3, I explain how we analysed the coherence of stories that users of the Interactive Storyteller created. Based on that study, we expanded our investigation and

¹ This chapter is based on Theune, Alofs, Linssen and Swartjes (2013) and Theune, Linssen and Alofs (2013).

developed an annotation scheme to classify the perspective of users, which I detail in Section 3.4. I close this chapter in Section 3.5 with the message that analysing the perspective of users on their interaction with the IST could inform the design of systems for interactive storytelling.

3.1 INTERACTION WITH INTERACTIVE STORYTELLING SYSTEMS

Children’s narratives have been the subject of many studies in the field of cognitive psychology, investigating the influence on narrative coherence of factors such as children’s age (Reese et al., 2011). In most of these studies, data were gathered by having children narrate the events shown in picture sequences. For example, Shapiro and Hudson (1991) conducted a study in which preschoolers were only able to construct coherent stories when they were supplied with pictures that corresponded to correctly formed story structures. First-graders showed to have a more expansive understanding of stories, as they were able to create more complex stories containing goals and plots. Trabasso et al. (1992) acknowledge Shapiro and Hudson’s findings. In their study, children as young as 4 years old used a goal plan to interpret pictorial representations of stories. This was supported by these children selectively encoding parts of the story that contributed to the plan. This process relies on intentionality: children are able to determine the intentions of story characters. The findings of the study of Low and Durkin (1998) on TV shows further support the conclusion of children being able to causally relate scenes. Children are highly organized when inferring relations between scenes, being able to determine goals, obstacles and repairs. Through imagining why a person would do something, children, just like older people, get a sense of the coherence of a story. Fisher (1989) states that *narrative coherence* is one of two prerequisites for *narrative rationality*, the other being *narrative fidelity*. The former requires a story to ‘hang together’ to make sense, while the latter entails that people only accept a story if it matches our experiences and beliefs. The work I discuss here is inspired by Fisher’s theory on narrative coherence and Trabasso et al.’s studies, in that we use similar methods to determine coherence. However, instead of studying narrations of given event sequences, we have looked at stories created by children using the IST, an interactive storytelling system.

In the past, several such interactive storytelling systems aimed at children have been developed. *FearNot!* is an educational system in which intelligent virtual characters autonomously act out bullying scenarios, with the children giving advice to the victim (Aylett, Louchart, Dias, Paiva & Vala, 2005, see Section 2.1.3.1). Preliminary evaluations showed promising results concerning the efficacy of this system, yet no studies investigating the children’s understanding of the stories were conducted. *Teatrix* is a tool for the development of narrative competence through collaborative story creation (Prada, Paiva, Machado & Gouveia, 2002). Multiple children can take on the roles of different characters and interact with computer-controlled characters in a virtual environment to create stories. They were also able to create their own scenes, deciding which props and characters to include. Prada et al. included a *Hot Seating* type of control which required children to explain their planned actions for a certain character, thus narrating

NARRATIVE
COHERENCE
NARRATIVE
RATIONALITY
NARRATIVE
FIDELITY

TEATRIX

how they felt the story would be causally connected (Paiva, Machado & Prada, 2001). A similar system was developed by Robertson and Despa (2006) in the form of *Ghostwriter*. This is a virtual role-playing environment for children, who interact with one computer-controlled character and two characters controlled by a human role-play leader. Stories children wrote after having played Ghostwriter were found to contain more complex character interaction than pre-game stories, indicating that the environment motivated the children to create characters with more depth (Robertson, 2000). Like the IST, these systems all take an emergent narrative approach and support children's story improvisation through roleplay in interaction with virtual characters. However, except for an analysis of the GhostWriter role play logs for indications of characters' personalities and moods (Robertson & Despa, 2006), we have found no analyses of the stories that have been created with these (or other) interactive storytelling systems.

Storytelling systems with tabletop interfaces, like the IST, include *Reactoon* (Alves, Lopes, Matos, Velho & Silva, 2010), *TellTable* (Helses, Cao, Lindley & Sellen, 2009) and *StoryTable* (Zancanaro et al., 2007). These systems intend to facilitate storytelling by children, allowing children to choose characters, create animations, and use physical objects in combination with a virtual environment. None of these systems incorporate intelligent agents as characters in the story. Moreover, in most cases, there appears to have been no evaluation of the stories created with these systems, with the exception of the StoryTable system. A comparison of stories created with or without the StoryTable revealed no significant differences in structure and cohesion (Zancanaro et al., 2007).

In the field of interactive digital storytelling, evaluation has received considerably less attention than system design and development. There is no generally accepted methodology for the empirical evaluation of interactive digital storytelling systems. Methods that have been used in user studies include questionnaires (Aylett et al., 2005; Vermeulen, Roth, Vorderer & Klimmt, 2010), interviews (Helses et al., 2009; Prada et al., 2002), analysis of system logs and created stories (Theune, Alofs et al., 2013; Zancanaro et al., 2007), and observations of user behaviour (Alofs et al., 2015; Helmes et al., 2009; Prada et al., 2002; Theune, Alofs et al., 2013). Aspects of the interactive narrative experience that are considered include the user's sense of presence and immersion, enjoyment, feelings of agency and control, believability of and identification with the characters, and system usability aspects. In this chapter, I discuss an approach that combines the analysis of created stories and user behaviour to provide additional insights into the process of creating stories using an interactive storytelling system.

3.2 STUDY SETUP

In their study on the Interactive Storyteller, Alofs et al. (2015) investigated to what extent this system supported social interaction, scrutinizing how it stimulated cooperation. Building on their work, we studied the coherence of stories the children co-created with the IST. First, I discuss the experimental setup Alofs et al. used before continuing to our analysis of the interactions in the next section.

GHOSTWRITER

REACTOON

TELLTABLE

STORYTABLE

Table 1.: The four pairs of children in the setup of Alofs, Theune and Swartjes (2015).

PAIR	PARTICIPANTS (AGE)	RELATIONSHIP	INTERACTION	
			SESSION 1	SESSION 2
1	Female (9) + male (9)	Friends	Touch-only	Tangibles
2	Female (10) + male (8)	Siblings	Tangibles	Touch-only
3	Male (8) + male (10)	Siblings	Touch-only	Tangibles
4	Female (10) + female (11)	Friends	Tangibles	Touch-only

Our analyses of the children’s interactions with the IST are fully based on the interactions recorded by Alofs et al. (2015), see Section 2.2.2. In their evaluation, four pairs of children interacted with the IST in two sessions, one with the use of tangible characters and one without, see Table 1. All participants were pairs of siblings or friends. One child controlled the character Red; the other controlled Granny. Wolf was always computer-controlled.

The experiment was originally set up to test, among other things, whether the use of tangibles in our system setup is preferred over a touch-only approach. Therefore, each pair of children carried out two play sessions, one with and one without tangibles (in a counterbalanced order). The results of this comparison are reported in Alofs et al. (2015); here, I only focus on the stories that were created by the children during the experiment. Because system version (tangible or touch-only) appeared to have no influence on the stories that were created, I ignore this factor in the remainder of this chapter.

The procedure of the experiment consisted of 8 steps as follows.

INTRODUCTION The facilitator explained to the children what the experiment would consist of and that their interactions would help the researchers improve the system.

EXPLANATION The children received an explanation about how to use the system for creating stories together, and on how they could interact with it. The facilitator did not give them a goal to attain with system, as our intention was to investigate spontaneous behaviour of the children.

ROLE ASSIGNMENT The IST allows users to take on any role in the story world. For this evaluation, Alofs et al. decided to let the Wolf be controlled by a Character agent, and Grandma and Red by the children, in all the play sessions. This way, the children would have a common antagonist. The children could chose themselves whether they would rather play as Grandma or as Red and situate themselves at the corresponding side of the table.

SESSION 1 The children interacted either with the version of the IST with tangibles or without.

QUESTIONNAIRE 1 The children answered questions on what they liked and disliked about the system and what they thought about the way in which they moved their characters

SESSION 2 The children interacted with the version of the IST they had not yet interacted with: with tangibles or without.

QUESTIONNAIRE 2 Again, the children answered the same questions as from Questionnaire 1.

INTERVIEW The facilitator conducted an informal interview with the children, focusing on the social interaction they experienced, which version of the system was more fun and which was easier to use, and whether they would like to play again with either of the versions. Additionally, the children answered whether they would prefer playing in couples or by themselves, and whether they had experience with multi-touch interfaces.

Each evaluation with the IST lasted for a total of about 45 minutes. The play sessions took between 9 and 14 minutes, with an average of 12 minutes of play time. This includes time spent on other matters such as resolving technical issues and asking questions. Video recordings captured the play sessions from both sides of the touch table to cover all angles. Additionally, they also captured the questionnaires and the final interview. All story actions taken through the Interactive Storyteller interface were logged, and the children's communication during the play sessions was transcribed. For the analyses of children's interaction I discuss below, I only consider the recordings of the actual play sessions, together with the characters' actions logged by the IST. In total, the video recordings added up to 146 minutes of play time.

3.3 COHERENT CO-CREATION

In this study, we explored the coherence of stories created by children with the Interactive Storyteller from three perspectives (Theune, Alofs et al., 2013). Firstly, we analysed which story components children selected (Section 3.3.1). Secondly, we investigated the causal connectedness of these components (Section 3.3.2). Thirdly, I discuss the connectedness from the children's points of view, based on video recordings (Section 3.3.3).

We analysed the stories generated in the experiment to see to what extent the children made use of the possibilities to create coherent stories offered by the Interactive Storyteller. Transcripts 4 and 5 show excerpts of two of the stories that were created. The fragments are shown as they were narrated in the story scrolls,² with each line describing the action of a character, controlled either by one of the children (playing the roles of Red and Granny) or an autonomous agent (Wolf). For convenience, the lines are numbered and the transcripts contains added information about the characters' initial locations between brackets. Because the actions are described one by one as they happen during the interaction, the

² The texts are translated from Dutch, see an example story in Dutch in Transcript 19 in Appendix A.

narrations are not very fluent, unlike the fabula-based narratives that could be generated by the VST after simulation is finished. However, in our story analysis we focus on the created story content and not how it is narrated.

3

1. [At Granny's house] Granny bakes the cheese cake.
2. Wolf sneaks to Red's house.
3. [At Red's house] 'Hello,' Red says to Wolf.
4. With a little bit of cyanide, Granny poisons the cheese cake.
5. 'Hello back to you,' Wolf says to Red.
6. Red does the dishes.
7. Granny shuffles to the beach.
8. 'Give me that!' says Wolf, and takes the birthday cake from Red.
9. 'Give that back, it is mine!' says Red, and takes the birthday cake back from Wolf.
10. Granny shuffles to Red's house.
11. 'Give that back!' says Wolf, and takes the birthday cake back from Red.
12. 'Oh Granny,' says Red, 'Wolf stole the birthday cake from me!'
13. 'Oh, I feel sorry for you,' Granny says to Red.
14. Wolf eats the birthday cake.
15. 'This is for you,' says Granny, and gives the cheese cake to Wolf.
16. Red bursts out laughing.

Transcript 4.: A story fragment from the first play session of Pair 4.

1. [At the lake] Red treads water for one minute.
2. [At the lake] Granny farts under water.
3. Red skips to the beach.
4. Granny bursts out in tears.
5. [At the lake] Wolf gets pretty hungry.
6. Red skips to Red's house.
7. Granny shuffles to the beach.
8. Red is angry at Granny.
9. Granny shuffles to Granny's house.
10. Red skips to the clearing in the forest.

Transcript 5.: A story fragment from the second play session of Pair 1.

3.3.1 Story components

First, we analyse the stories created by the children in terms of their components: the character actions the children selected during their interactive storytelling sessions, as captured in the system logs.

Previous studies of story coherence have looked at the presence of narrative components such as setting information, character descriptions, actions, dialogue, internal responses, obstacles and repairs (Low & Durkin, 1998; Shapiro & Hudson, 1991). Of these component types, only three are available for selection by the players in the Interactive Storyteller: dialogue, internal responses (emotions), and physical actions. Interactions with the IST always started with the system narrating the following introduction in the story scrolls: ‘Once upon a time, there was a little girl wearing a little red riding hood. She wanted to bring a birthday cake to her grandmother. . .’ While this text was meant to incite the player controlling Red to bring Grandma the cake, the system by no means forced this goal upon the player. The IST contains no mechanism for determining the goals of players, which results in a lack of information on players’ goals in our story logs. Therefore, we attempted to analyse the player goals by looking at the story components players used. Players can choose from three types of components in the IST: dialogue, internal responses (emotions), and physical actions. For physical actions, we distinguish between *causal chain* and *dead-end* events, based on Trabasso, Secco and Van den Broek (1984). Causal chain events are causally connected to other events in a goal-action-outcome sequence, while dead-end events have no follow-up and do not lead to goal satisfaction or failure. In Trabasso’s causal network model of stories, causal chain events are causally connected to other events in a goal-action-outcome sequence, while dead-end events have no follow-up and do not lead to goal satisfaction or failure (Trabasso et al., 1984). While Trabasso’s division between causal chain and dead-end events is based on a causal network analysis of actual stories, our action type classification is an *a priori* one, based on our estimation of the likelihood that actions of a certain type will be causally connected to other actions. It does not tell us whether the actions will actually end up as part of a causal chain or a dead ends in a story. We added the class of move actions as a third class, separate from the first two types of events, as it is hard to predict whether moving characters around in the IST is part of a causal chain. Move actions can go either way, and in that respect they are somewhere in between causal chain and dead-end actions. A summary of the five types of story components follows below.

CAUSAL CHAIN ACTIONS Actions which entail preconditions for other actions, such as baking and giving away cakes. These have the highest potential for creating coherent stories, as they may be part of a goal-directed action sequence: characters can plan a series of such actions to try to achieve some goal. For example, a plan to poison Wolf might involve first baking a cake, then poisoning it with cyanide, and finally giving it to Wolf (expecting him to eat it). Each action in such a sequence is a prerequisite for the next.

DEAD-END ACTIONS Actions which do not have causal links to any other actions, for example, diving into the lake. They are the opposite of causal chain actions in that they have the least potential for contributing to story coherence. They do not change the story world in our implementation and thus cannot serve as prerequisites for any other actions: no new actions become available after a dead-end action has been carried out. This makes these actions dead ends by definition.

MOVE ACTIONS Actions used to go to a certain location for various reasons. Players might choose to move to establish a precondition of another action, for example, a character can only give a cake to another character if both are at the same location. To achieve this, a sequence of move actions may be necessary. Fragment 1 exemplifies this: Granny goes to Red's house via the beach as part of her plan to poison Wolf, who is at Red's house. On the other hand, players may also move between locations without a clear goal in mind. They may be carried out more or less at random, as in Transcript 5, where Red is skipping around without any clear in-story purpose.

EMOTIONS Actions used to express internal states. In the IST, these are limited to anger, happiness (laughing) and sadness (crying). Players can use such actions to adopt a goal, for example, when Red gets angry at Wolf, she may adopt the goal of poisoning him. Crying may also indicate story coherence, for example, when Red starts crying because Wolf ate the cake meant for Grandma.

DIALOGUE Actions used to discuss events or to greet each other. In the IST, such interactions are linear: every utterance of a character only enables one causally connected response. One dialogue action often triggers another, leading to brief exchanges like the one in Transcript 4, where Red complains to Granny about Wolf having taken her cake, and Granny reacts by expressing her sympathy. The boundaries between dialogue and emotion are not very clear-cut: expressing sympathy through dialogue could also be seen as an indication of emotion (feeling sorry for the other), and the same could be said of saying thanks as an indication of gratefulness. However, we feel the dialogue aspect is more prominent in these cases.³

Table 2 shows the frequency of the different story components across all the play sessions of the four pairs of children. With a total of 347 character actions over all play sessions, the children took a joint average of 43 actions per play session. Wolf, the character controlled by an Actor agent, only took 69 actions across all sessions (17% of the total). The reason for this relatively low percentage is that Wolf only carried out goal-directed actions (in his case, chasing after cakes to eat) and frequently skipped a turn when unable to make a plan to achieve his goal of eating cake.⁴

³ As can be seen in Transcript 4, some character dialogue is included in the narration of GiveTo and TakeFrom actions. However, these actions are not presented as dialogue actions to the players; the dialogue is only added to spice up the narration of the physical actions.

⁴ In the RED domain only Granny and Red can bake cakes, so Wolf could not bake his own.

Table 2.: The frequency of the five types of story components in the experiment.

COMPONENT TYPE	RED ACTIONS	AMOUNT	FREQ. (%)
Causal chain action	Bake(Cake), Poison(Cake), GiveTo(Cake, Character), TakeFrom(Cake, Character), Eat(Cake)	71	20
Dead-end action	WashDishes, DustCabinet, DiveIntoLake, TreadWater, FartUnderWater, EnjoyTheSun, RollInSand, TakeNap, BuildSandCastle, WatchBirds, WatchClouds, LeaveAlone(Character)	71	20
Move action	MoveFromTo(CurrentLocation, NewLocation)	111	32
Emotion	Laugh, Cry, BeAngryAt(Character)	67	20
Dialogue	Greet(Character), TellAboutCakeTaken(Character), ExpressSympathyFor(Character), Thank(Character), AskWhatToDo(Character), TellAboutPlan(Character)	27	8
TOTAL		347	100

The results in Table 2 suggest a preference of the children to carry out actions that were likely to be part of a causal chain. While four of the five locations made at least two dead end actions possible to be carried out, the children selected just as many causal chain actions as dead-end actions. However, this suggestion does not mean that the selected actions were followed up with a complete causal chain, leading up to fulfilling a goal. Therefore, we investigated the causal connectedness of the story components, which I discuss in the following section.

3.3.2 Causal connectedness

To see whether the selected story components were really part of causal chains or not, we examined the causal relations between them. To this end we looked at the stories produced by the children (as shown in Transcripts 4 and 5), and used common sense reasoning to find logical connections between the story components. The VST, the non-interactive version of our system, logs all causal connections between character actions in the fabula model. However, in the Interactive Storyteller these causal connections are only recorded for computer-controlled characters. For player-controlled characters, the system only records which actions they took, because it has no knowledge of their plans and motivations. Therefore, we established a method to manually determine the connectedness of story actions.

To count as causally connected, story components either had to be part of a goal-directed causal chain in the sense of Trabasso et al. (1984), thus providing global coherence, or be connected to an immediate cause or consequence, thus

providing local coherence (Trabasso et al., 1992). Simple enablement relations were only counted if they were clearly part of a plan, for example, a move to some location enabling a character to do something at that location. Unlike Low and Durkin (1998), in our analysis we did not limit ourselves to determining causal relationships between adjacent story components. In the Interactive Storyteller, characters take turns to carry out actions in the story world. This means that in the narration, components of an individual character's plan are usually not adjacent but intersected by other characters' actions.

As an illustration, consider the two story fragments in Transcripts 4 and 5. Transcript 4 holds a clear example of a goal-action-outcome sequence with Granny carrying out a plan to poison Wolf, baking a cheese cake as the first step (line 1), poisoning it (4), and going to Wolf (7, 10) to give it to him (15).⁵ Red, meanwhile, carries out more locally connected actions: she reacts to Wolf's arrival at her house by greeting him (3), and to his repeatedly taking her birthday cake. The first time, she takes it back (9) and the second time she complains to Granny (12). In her turn, Granny briefly interrupts her poisoning mission to sympathize with Red (13). Finally, Red reacts with laughter (15) to Granny giving the poisoned cake to Wolf. Overall, Transcript 4 is quite coherent. The only action that is not causally connected is Red doing the dishes (6), a dead-end action with no apparent cause or purpose within the story.

Transcript 5, on the other hand, is incoherent: the characters seem to be carrying out actions at random, and when they display emotions it is not clear what these are in response to. I annotated all stories produced in the experiment together with another researcher. For each character action, dialogue, and emotion, we determined whether it was causally connected to another story component; in other words, whether that component was coherent or not. Annotation took place in three rounds. In the first round, half of the stories were annotated by both annotators. The inter-annotator agreement for this first round was fairly low (Cohen's $\kappa = 0.60$). After discussing and fine-tuning our definition of connectedness, the remaining stories were annotated, this time achieving a satisfactory level of agreement (Cohen's $\kappa = 0.73$). Finally, in the third round all differences between the annotators were resolved by discussion. The results are given in Tables 3 and 4.

Table 3 shows, per story component type, how many instances of that type were causally connected to some other component, or not. It also shows how many story components were causally connected overall, regardless of their type. Taking together all components of all stories, 40% of these components were causally connected (coherent) while 60% were not. Table 4 shows the proportion of all connected or disconnected story components that belonged to each of the component types, thus showing the relative contribution of all types to the overall coherence of the created stories.

Unsurprisingly, most causal chain actions turned out to be causally connected, and causal chain actions contributed the most to story coherence, with the opposite holding for dead-end actions. We found two dead-end actions that actually seemed to have a causal connection to another story component: one was

5 The eventual outcome, Wolf eating the poisoned cake, is not included in the excerpt.

Table 3.: Causal connectedness of story components per type.

COMPONENT TYPE	CAUSALLY CONNECTED (%)	NOT CAUSALLY CONNECTED (%)
Causal chain action	75	25
Dead-end action	3	97
Move action	32	68
Emotion	34	66
Dialogue	93	7
All types	40	60

Table 4.: Contribution to story coherence of each component type.

ACTION TYPE	CAUSAL (%)	CHAIN	DEAD-END (%)	MOVE (%)	EMOTION (%)	DIALOGUE (%)
Coherence	38		1	25	17	18
Incoherence	9		33	36	21	1

Granny rolling in the sand at the beach, triggering an emotional response by Red (laughing), and the other was Red leaving Wolf alone after having given him a cake to eat. Only 36% of the move actions turned out to be causally connected, either because they were part of a plan or because they triggered an immediate reaction such as a greeting. Due to their frequency, moves still provided the second largest contribution to story coherence. Across all stories, we counted 18 goal plans (goal-directed action sequences). These causal chains had an average length of four story components. Of the emotions, only 34% were annotated as being causally connected, usually because they were clearly a response to another story component. In contrast, dialogue connectedness was 93%. Dialogue thus provided a relatively large contribution to story coherence in spite of being the least frequently selected component type. The most frequent dialogue actions were exchanges of greetings, which provided local rather than global coherence, similar to emotions. A mixed picture emerges from this analysis. On the one hand, a large portion of the created stories was found to be coherent in terms of the number of causally connected components, displaying both global and local coherence. On the other hand, an even larger portion of the stories was incoherent, consisting of seemingly unconnected story components. However, this analysis presents a somewhat misleading image of the stories' coherence, because not all that makes the stories coherent has been captured in the logs of the IST. Specifically, what is missing are the verbal and non-verbal contributions the children made to the story while they were interacting with the system and with each other. It is this additional layer in the interaction that allows us to get a better understanding of the children's intentions, as they often went meta by discussing various facets of their interaction from an out-of-character point of view. In the following section, I address how we analysed the children's communication.

3.3.3 Children's communication

The emergent narrative approach to interactive storytelling is closely related to improvisation and children's dramatic play. As argued by Sawyer (2003), transforming an improvisation into a fixed text removes its most essential aspect, which is that it is a social, collaborative process, in the form of a dialogue between players. Therefore, it is important not to overlook the children's communication when analysing the stories they created. In the Interactive Storyteller, what happens in the story world is automatically narrated through synthetic speech and shown in the graphical interface. Thus, in principle, stories can be created without any spoken communication between the human players. In practice, we found that most of the children did spontaneously communicate with each other during the experiment, thereby adding an extra layer to the emergent story. We inspected the transcripts of children's communication to get some insight in their goals and motivations for selecting certain story components. Many of the selected actions went undiscussed by the children, with one pair of children being too shy to say much at all. As a result, the following observations are anecdotal in nature. First, we examined the communication transcripts to see if they provided any evidence for or against the in-story goals and motivations we had inferred as part of our story analysis in the previous section. For certain character actions, this inspection revealed external (out-of-story) motivations or unexpected in-story motivations were revealed. These utterances revealed that the children approached the interaction from a meta-perspective. For example, one child announced '*I am going to bake, that is fun!*' indicating that this causal chain action was not initially selected as part of a plan, while another child explained '*I am giving the [poisoned] chocolate cake to Wolf. He didn't say anything, so I had to poison him.*' Another examples is the following exchange:

Action [by Child 1]: Red eats the birthday cake.

Child 2: 'Actually, Red should have given the birthday cake to Granny.'

Child 1: 'I already ate it. Otherwise the wolf would have eaten it.'

Transcript 6.: Example exchanges of children when interacting with the IST (Pair 2, session 2).

Not all goals mentioned by the children were actually followed up by actions. An example is a child (controlling Red) exclaiming '*I want to kill Granny*' after having been angered by the child playing Granny. However, after an alternative proposal by the other child, it was jointly decided to poison Wolf instead.

The presence of goal plans was often revealed through the children discussing what to do next. Examples are the following.

Child 1: 'Bake another cake.'

Child 2: 'That can only be done in Granny's house.'

Action [by Child 2]: Granny shuffles to Granny's house.

Action [by Wolf]: 'Hello!' Wolf says to Red.

Child 1: 'You have to give him the cake!'

Child 2: 'It has to be poisoned first.'

Child 1: (...) 'I will do that, alright?'

Child 2: 'OK.'

Action [by Child 2]: Granny poisons the chocolate cake.

Child 1: 'I am going to tell Wolf that you took my chocolate cake!'

Child 2: 'No, I am poisoning it.'

Child 1: 'Yes, but then I'll give it to Wolf.'

Transcript 7.: Example exchanges of children when interacting with the IST (Pair 3, session 2; Pair 4, session 2; and Pair 1, session 1, respectively).

When the children explicitly mentioned making plans for the next actions of their characters, this usually happened late in the first interaction session, or in the second, after the children had gained some insight in the available actions in the story world and the causal connections between them. Overall, it seems that many causal chain actions were initially selected just like all other actions, as part of an exploration of the story world, to see what would happen, or simply because they were thought to be 'fun'. However, most children quickly discovered that these actions could be used to achieve goals such as giving away cakes and poisoning other characters (usually Wolf, but sometimes also their playmate's character), and then they invariably started making plans to do so. Those plans were sometimes abandoned prematurely, or the causal chains were disrupted by other actions, for example if the children were distracted from their plan by location-specific dead-end actions. The children did not appear to perceive this as incoherent. In line with Sawyer's observations of children's dramatic role play (Sawyer, 2003), they were caught up in the moment-to-moment contingency of the improvisation process and did not seem overly concerned with the global coherence of the emerging story. This is similar to adult role players, who focus on a local rather than a global story level (Louchart & Aylett, 2003). We noted one possible exception, where the children did not want to end their play session until Wolf had eaten the cake they had poisoned for him. However, it seems likely that these children simply wanted to see the achievement of their goal, and did not consciously aim for narrative closure. The children's dialogue also showed that several actions that seemed incoherent within the story narration were not incoherent at all in the context of the improvisation. For example, it turns out that Granny's crying in Transcript 4 did in fact have a causal connection with the preceding action (Red skipping to the beach), as the child controlling Granny provided the in-character motivation 'Because you are leaving me!' Another seemingly incoherent emotion in the same fragment is Red getting angry at Granny for no clear reason. Inspection of the

children's communication transcripts revealed that Red's emotion was actually a reaction to an event outside the story:

Child 2 [holding tangibles up to camera]: 'Isn't this a nice Granny? Isn't this an ugly Red Riding Hood?'

Action [by Child 1]: Red is angry at Granny.

Transcript 8.: Example exchange of children when interacting with the IST
(Pair 1, session 2).

Another example of a child providing a motivation for a seemingly incoherent emotion is the following:

Action [by Child 2]: Granny bakes the cheese cake.

Action [by Child 1]: Red bursts out in tears.

Child 1: 'Why can't I bake the cake?'

Transcript 9.: Example exchange of children when interacting with the IST
(Pair 3, session 2).

In this fragment it is not entirely clear whether the child was speaking in character or out of character, but it was most probably the latter. This means that like in Transcript 8, Child 1 used the character to express his own emotion. These examples show that the children did not always distinguish between what was part of the 'story proper' and what was not; they were busy playing, not necessarily creating stories. Through their play communication, the children did not only motivate and negotiate the character actions, but also enriched the emerging story in various ways. In a few cases, instead of selecting dialogue actions for their characters through the system, the children carried out the character dialogue in person, for example thanking the other or expressing sympathy for the loss of a cake. The fact that dialogue could be more easily expressed via direct communication than via the system may explain the relatively low frequency of dialogue among the story components. In most cases, the children's play-acting did not replace the system's narrations, but augmented them. The children frequently played out what happened in the story world, expressing their character's emotions through sounds and facial expressions, and miming character actions such as shuffling, diving and eating, often while repeating the system's narrations. In other cases, the children brought in new elements that were outside the possibilities of the system. They added their own narrations ('Much later... Wolf is still trying to think of a plan.'), expressed character emotions that were not available in the system ('Ooh, I'm scared!'), invented imaginary props (a cake-mould to bake a cake), and even made up entirely new actions (Red kicking Granny out of the water at the lake). The observations presented in this section show that the actions that were logged and narrated by the Interactive Storyteller do not tell the whole story. The children's communication is an integral part of the improvised narratives, supplying causal connections that may not be obvious at first sight, and enriching the story in various ways.

3.3.4 Discussion

When looking at the components of the stories created in our experiment, we see that the children mostly used actions that had a strong potential for being causally connected. We call these *causal chain actions*. They used fewer actions with little such potential, which we call *dead-end actions*, than might have been expected based on the availability of these action types. Inspection of the causal links between these components, as inferred from the narrations, showed that 40% of all story components were causally connected. Moreover, when seen in the context of the children's play most stories turned out to be more coherent than they seemed to be when judging only from what was logged by the system. The children's system-external communication confirmed that they were acting in a goal-directed fashion at least part of the time. It is through these meta-communicative acts that the children were able to collaborate and plan towards desired outcomes of the stories. In Section 3.4, I delve deeper into the types of information the children shared while interacting with the IST.

At the same time, it is clear that the children did not use all available opportunities for achieving story coherence, in the sense of maximizing the causal connections between the story events. Unlike the computer-controlled character Wolf, they did not always plan ahead, and if they did, they did not always follow through with their plans, allowing causal chains to be disrupted by unrelated actions. One possible reason for this is that at first, the children were not yet aware of all available actions and their effects. They only gradually discovered the things they could do within the story world, and this prevented them from exploiting all opportunities for goal-directed action right away. Our data support this explanation: 4 goal plans were carried out in the first play sessions, against 14 in the second. The proportion of causally connected story components was 36% for the first sessions, and 43% for the second.

Another possible reason for the suboptimal coherence was that, although the children had been told they could use the system to create stories, they had not been given the explicit goal to do so. Especially in the first sessions, much of the time the children were simply exploring the story world and trying out different actions without any concern for creating a coherent narrative. In general, the children's story making was done in a moment-to-moment fashion, with the children reacting to each other's and Wolf's actions as they happened. This resulted in the children achieving local rather than global coherence. According to Sawyer (2003), this is typical for children's dramatic role play: children who are engaged in such play do not attempt to create an overall coherent narrative, but instead react to each other turn by turn, which leads to 'pockets of local coherence' rather than global coherence in Trabasso's sense.

Overall, the children in this experiment were not overtly concerned with producing coherent narratives. Nevertheless, they spontaneously carried out many character actions that led to narrative coherence. Their motivation for creating globally coherent action sequences may have been simply that they thought, for example, poisoning another character would be a fun thing to do.

In these cases, the players' goal of having a satisfying experience happened to coincide with the requirements of a coherent narrative.

Similar to roleplay, in emergent narrative the players' experience is more important than the resulting story as seen from an observer's perspective. In other words, the process is more important than the outcome (Louchart & Aylett, 2003). To support this process, all kinds of actions are needed. Causal chain actions allow the players to achieve goals; emotions and dialogue allow them to express themselves and communicate with computer-controlled characters; and dead-end actions can simply provide fun things to do in the story world. No strong conclusions can be drawn from this small-scale experiment as to which types of actions are most important to achieve both player satisfaction and story coherence. However, I hypothesize that providing more causal chain actions, as suggested by Swartjes and Theune (2009a), may be most beneficial for this purpose, as it will allow the users to have a major influence on the course of the story (global agency (Mateas & Stern, 2005b)), and thus presumably increase their enjoyment. As a side-effect, the generated narratives will also be more coherent.

To test this hypothesis, larger and more focused user experiments will be necessary. In such experiments, the players should first be given the opportunity to thoroughly explore the story world and acquaint themselves with the system, to make sure that any lack of coherence in the created stories cannot be attributed to the players' lack of awareness of the possibilities. In addition, it might be useful if players could not only select character actions but also character goals through the system's interface. Finally, the players could be asked to think aloud to provide more insight in their reasoning and motivations than could be gained from their spontaneous communication. A drawback of such an instruction is that it is likely to interfere with the children's improvisation. A further alternative is to use a 'talk aloud' approach where children are instructed to talk about what they are doing instead of what they are thinking (Donker & Reitsma, 2004). This is supposed to be easier for children, but may still suffer the same problem of interfering with improvisation. On the other hand, unintrusive methods such as retrospection and post-task interviews are expected to be less informative.

Future experiments should be carried out not only with children, but also with adult players, using an appropriate story domain. Previous studies have shown that at age 8, children 'seem to be on the verge of a more sophisticated understanding of goals' (Lynch & van den Broek, 2007, p. 327) and at age 9, they perform at a similar level as adults with regard to narrating stories in terms of goal plans (Trabasso et al., 1992). Given their age group, the children who participated in this experiment should therefore have been capable of generating equally coherent stories as adults. However, children tend to be more playful and have a shorter attention span than adults, and this may have influenced the results. Although these findings are not generalizable to adults, I expect that adult players will have similar preferences as children in regard to story component types and goal-directed planning.

3.4 PERCEPTION OF INTERACTION

In the previous section, I described our efforts to evaluate the coherency of children's stories created with the Interactive Storyteller. A key element in this investigation was the verbal communication between the children outside of the IST. In this section, I expand on the analysis of their communicative acts, determining which perspectives they took while interacting with the system and with each other. In an effort to obtain quantitative results, not just anecdotal evidence, we designed a multi-layered scheme for annotating the children's utterances. The scheme is based on the notion of *framing* as used in research on role-playing games (Fine, 2002) and children's pretend play (Sawyer, 2003). A central notion for both is that of framing, a term that was first introduced by Bateson (1972), and greatly expanded on by Goffman (1974). We have used this scheme to annotate the transcripts of eight interaction sessions with our storytelling system, involving four pairs of 8–11 year old children. Below, I present and discuss the results.

FRAMING

3

We are not aware of any user studies that systematically analyse the communication of people interacting with a storytelling system. However, users' communication can be a rich source of information about, among other things, their level of identification with the characters, what they think of the story, and their understanding of the system's interface. In the study discussed in the previous section, the following fragment was found to contain an incoherent sequence of actions when only looking at the story logs.

Action [by Child 1]: Red skips to the beach.

Action [by Child 2]: Granny bursts out in tears.

Child 2 [in character, addressing Child 1]: 'Because you are leaving me!'

Transcript 10.: Example exchange of children when interacting with the IST.

In this transcript, Child 2 is role-playing as his character, Granny. He addresses the other child (Red) in character, providing a motivation for Granny's emotional outburst. He is actively taking part in the story creation process, identifying himself with the character he is playing. In this section, I address how this additional information offers insight into the intentions of players.

Not all of the children's communication demonstrates the same level of immersion in the story being created. In the next transcript we see Child 6 taking a game-playing rather than a role-playing approach to the storytelling activity:

Table 5.: Sawyer’s annotation scheme for the frame dimension (adapted from Sawyer (2003)).

LEVEL	FRAME	EXAMPLE COMMUNICATION
1	In-frame	Imitating voices, referring only to in-frame objects or events. For example, ‘Watch out for the earthquake!’
2	Blended frame (in-character)	Speaking in-character, but referring to the world outside play. For example, in a character’s voice: ‘I need more blocks!’
3	Blended frame (out-of-character)	Speaking out-of-character about the play itself. For example, ‘Let’s say my guy was killed in the earthquake.’
4	Out-of-frame	Speaking out-of-character, not about the world in the play frame. For example, ‘Let’s play house now.’

Action [by Child 5]: Red eats the birthday cake.
Action [by Child 6]: Granny shuffles to the beach.
Child 6 [addressing Child 5]: ‘Then it’s your turn.’
Action [by Child 5]: Red does the dishes.
Child 6: ‘And now Granny.’ [selects an action] ‘That one, I think.’

Transcript 11.: Example exchange of children when interacting with the IST.

In this section, I first discuss the works of Sawyer (2003) and Fine (2002) (Section 3.4.1) before explaining our P×R annotation scheme (Section 3.4.2). I discuss our analysis of children’s communication during interaction with the Interactive Storyteller in Section 3.4.3. In Section 3.4.4, I discuss the validity of the annotation scheme.

3.4.1 *Frames of play*

Given that our approach to interactive storytelling is closely related to children’s improvised dramatic play, we first looked to work in this field for inspiration. Sawyer (2003) categorizes meta-communication in pretend play along two dimensions: *frame* and *explicitness*, distinguishing four levels within each dimension.

Explicitness only concerns children’s proposals to modify the play frame by changing the storyworld. In the Interactive Storyteller this is done via the system, so I do not discuss this dimension further here, focusing only on frame. The communication’s frame can range from in-frame (in-character) to out-of-frame, with two blended levels in between, as shown in Table 5. In the blended frames, reference is made to objects or events from the other frame.

Sawyer’s notion that in-frame and out-of-frame communication are not wholly separated realms, but that they can be blended, is a useful starting point. However,

we feel that Sawyer's scheme is insufficient for our purpose. In the Interactive Storyteller, unlike dramatic play, the children do not only interact with each other directly but also via the system, which provides a shared visual representation of a story world and external characters that can be controlled. This means there is a division between the story that is created during play and the play itself. Sawyer's annotation scheme for frames cannot account for this distinction, which means that a more fine-grained annotation scheme is called for.

Whereas Sawyer only distinguishes two basic frames of play and non-play, with blending between them, Fine (2002) distinguishes three frames in his study of fantasy role-playing games. Each of the three frames is linked to a different side of the role-player's identity:

1. The 'primary framework' of *reality*, with participants as persons in the real world;
2. The framework of the *game*, with participants as players manipulating their characters within the rules and constraints of the game;
3. The framework of the *fantasy*, with participants 'being' the characters that they play.

Interactive storytelling with the Interactive Storyteller is similar to fantasy role-playing, in that players assume the roles of characters in an imaginary world of which the setting is given, and do so within a (non-competitive) game context with certain rules and procedures. This makes Fine's frame structure suitable for the IST. The interaction fragments from the introduction illustrate two of Fine's frames: the child as character in Fragment 1 and as player in Fragment 2. An example of a child's utterance in the primary framework is '*He's still filming us!*' referring to the video recordings that were made during the user experiments (see Section 3.4.3.1).

Fine observes that during the game, the players are constantly oscillating between frames, and that participants are 'able to operate on several levels nearly simultaneously' (Fine, 2002, p. 240). However, unlike Sawyer, he does not explicitly distinguish blended frames as separate levels.

3.4.2 The $P \times R$ annotation scheme

In our opinion, both Sawyer's and Fine's frameworks are not fine-grained enough to capture the different types of communication that occur during interaction with our storytelling system and, presumably, other systems. However, we feel that the two can complement each other to allow for sufficiently rich annotation. Sawyer's scheme incorporates the possibility of blending, which is necessary because communication can be in one frame while referring to another frame. Fine's scheme includes a game frame, which is necessary to allow for annotation of communication about game play. Therefore, we combined Sawyer's and Fine's ideas to create a new annotation scheme that we call $P \times R$. The new scheme includes Fine's three frames but also allows for blending between frames. Our annotation scheme (see Table 6) is based on the notion that a user of an interactive storytelling system can take on different *perspectives* (P) toward different frames of

$P \times R$
ANNOTATION
SCHEME

Table 6.: The P×R annotation scheme, including shorthand notation for the different communication categories.

		REFERENCE		
		Story	Game	Reality
PERSPECTIVE	Character	CS: In-character utterances and imitations	CG: In-character references to game elements	CR: In-character references to events or objects outside play
	Player	PLS: Action suggestions and proposals referring to the story	PLG: Communication about game aspects	PLR: Including real-life events or objects in the game frame
	Person	PES: Observations about events that happened in the story	PEG: Observations about the interface, opinions about the game	PER: Communication about events or objects outside play

reference (R). In other words, the scheme indicates which identity a speaker takes on (perspective) and what she speaks about (reference), yielding a total of nine different classes. The abbreviations for these classes are constructed by taking the first one or two letters from the perspective dimension and appending the first letter of the reference dimension to them.

We want to be able to annotate how a user relates to the story: as a character in the story, as a player controlling a character in the story, or as an observer of the story. These classes of utterances are captured by the CS, PLS and PES categories respectively. We also want to distinguish whether the user is referring to the story, the game or the reality. This is an extension of Sawyer’s distinction between in-frame and out-of-frame references, with Fine’s game frame as an added out-of-character level.

By combining the two dimensions of perspective and reference, we broaden the scope of Fine’s categorisation, which would only allow for annotation of the classes on the diagonal of our scheme (CS, PLG and PER). A mapping of Sawyer’s scheme to our own yields the following correspondences of our categories to his frame levels is shown in Table 7. This illustrates that our scheme allows for distinctions that Sawyer’s cannot make.

As shown above, the P×R annotation scheme offers several refinements of Sawyer’s and Fine’s schemes. There are however some classes that are less intuitive, namely the CG, CR and PLR classes. A fictional example of a CG utterance would be a child speaking as Red about the game saying that ‘It’s too bad I have to wait for Granny before I can do anything,’ indicating in-character that she would have to wait for her turn in the game. Granny might also exclaim ‘Oh, Red, we are being filmed!’ This would be annotated as a CR utterance, as it is an in-character reference to an event taking place outside the story and game frames. Lastly, a PLR utterance

Table 7.: Mapping of Sawyer's scheme to the P×R scheme. The shades of grey indicate Sawyer's levels: level 1, level 2, level 3, and level 4.

		REFERENCE		
PERSPECTIVE		Story	Game	Reality
	Character	CS	CG	CR
	Player	PLS	PLG	PLR
	Person	PES	PEG	PER

is one in which a user speaks as a player about the reality frame, for example, a user tries to incorporate an object or event from the reality frame (outside the game and story) into the game. This would be the case when a user says that she can only go swimming at the beach when it is sunny outside; she introduces a game rule based on the reality frame.

3.4.3 Analysis of children's communication

In this section I describe the communication data to which we applied the P×R annotation scheme. I also describe the annotation process, and present and discuss the results of this study.

3.4.3.1 Data and annotation

The communication data were collected in a small-scale experiment, in which four pairs of 8–11 year old children played with the Interactive Storyteller (Alofs et al., 2015). Section 3.2 contains the details of this experiment. For this study, we segmented the transcripts of the interaction into utterances based on speaker turns, bounded by silence or speech by the other child or the system. When a speaker turn consisted of multiple utterances that should each be assigned a different label, the turn was split. This was done in only four instances; all other speaker turns were treated as one utterance.

All utterances were annotated by the first two authors of Theune, Linssen and Alofs (2013), using the coding scheme from Table 6. First we carried out two practice rounds, annotating one and two play sessions respectively, and comparing and discussing the results to come to a shared understanding of the annotation scheme. The remaining five sessions were annotated independently (Cohen's $\kappa = 0.74$, indicating substantial agreement). As a final step, all differences between the annotators were resolved by discussion.

In total, 640 utterances were annotated. This number excludes those cases where we could not make out what a child was saying.

Table 8.: The annotation results of the children’s communication in their first sessions with the interactive storytelling system.

		REFERENCE			Total (%)
		Story (%)	Game (%)	Reality (%)	
PERSPECTIVE	Character	6	0	0	6
	Player	25	32	0	57
	Person	13	23	1	37
	Total	58	38	4	100

Table 9.: The annotation results of the children’s communication in their second session with the interactive storytelling system.

		REFERENCE			Total (%)
		Story (%)	Game (%)	Reality (%)	
PERSPECTIVE	Character	7	0	0	7
	Player	36	22	0	58
	Person	15	16	4	35
	Total	58	38	4	100

3.4.3.2 Results and discussion

Tables 8 and 9 show the results of the annotation. They are separated by session, because the children’s first and second interaction sessions had different distributions of utterance types.

We found some individual differences in communication between the pairs of children: Pair 1 talked more in character, Pair 2 talked more from a ‘person’ perspective, and Pair 4 communicated more in general than the other pairs. However, the overall image was similar for all pairs.

In their first sessions with the system, the children communicated more about the game, while in their second sessions they communicated more about the story. In both cases, they were mostly speaking from a player perspective (57% and 58%, respectively). These findings may be explained by the fact that in the second sessions, the children were more familiar with the system’s interface and its rules and constraints. This led to fewer discussions about, for example, turn taking or which locations were accessible to the characters (both annotated as PLG) and fewer comments on the quirks of the interface (annotated as PEG). In addition, by the second session they had discovered more of the things the characters could do in the story world, resulting in more mutual action suggestions and discussing plans for their characters (annotated as PLS). An example is the following exchange:

Action [by Wolf]: Hello, says Wolf to Red.

Child 7: 'You have to give him the cake!'

Child 8: 'It has to be poisoned first.'

Child 7: '(...) I will do that, alright?'

Transcript 12.: Example exchange of children when interacting with the IST.

That such PLS exchanges occurred more often in session 2 is in line with the findings in the previous study (see Section 3.3). In that study, we found that the stories created in the second sessions contained longer causal chains of events, thus showing more coherence. Another, very noticeable, difference between the first and second sessions is that the children generally talked more in their second session. This is probably because by the second session, the children had become more at ease with the experimental setting and thus talked more freely.

We also compared the utterance type distributions of interactions with the tangible or the touch-only version of the system. Here, the only remarkable difference in the children's communication was that the children more often took on a 'person' perspective with the tangible version than with the touch-only version. A possible explanation for this is that when the characters were personified as tangible objects, the children were more likely to see them as separate entities, and therefore identified less with them. On the other hand, exchanges such as the following (coded as PEG) show that the children did see the tangibles as representations of themselves:

Child 1: 'Oh, you are as fat as Granny! Granny is fat.'

Child 2: '(...) In reality, I'm really thin!'

Transcript 13.: Example exchange of children when interacting with the IST.

As can be seen in Tables 8 and 9, relatively little of the communication was carried out in character (CS; 6% on average). This is not surprising, because the design of the system did not particularly encourage it. As mentioned in Section 3.2, all character actions are narrated through speech by the system, in principle making it unnecessary for the children to add their own dialogue. Nevertheless, the children did speak in character to express story elements that were not available in the system, such as certain character emotions ('Oooh, I'm scared!') and goals ('Especially for you, Red!' when baking a cake). In-character communication also frequently involved miming character actions such as shuffling, diving and eating, and expressing the characters' emotions through sounds and facial expressions.

We noted that in most PES utterances the children talked about their characters in the third person, while in most PLS utterances they used first person pronouns ('I am going to bake the apple pie.'). Pronoun use is an interesting aspect to investigate further, as the users' choice of pronouns indicates how strongly they identify with their character, and, in general, to what extent they are engrossed in the activity (Stromberg, 2000).

3.4.4 *Validity of the P×R scheme*

In the previous section I showed that applying the P×R scheme to the children's communication data provided us with several useful insights. In this section, I discuss the scheme's reliability and examine its validity by viewing the results through the lens of the other frameworks. I also mention some points for improvement of the annotation scheme.

For the five sessions that were annotated independently (see the previous section), we computed an inter-annotator agreement of 0.74 (Cohen's κ), which indicates good agreement and thus reliability of the coding scheme. The only main disagreements between the annotators concerned utterances in which the players commented on the actions of Wolf, the computer-controlled character. One annotator coded all of these utterances as PES, seeing them as person-perspective observations on the story, specifically, Wolf's actions within the story. The other annotator coded several of these utterances as PLS, because they referred to Wolf as a fellow player of the game rather than as a character in the story. This is illustrated by the following transcript:

Child 4: 'Doesn't Wolf have any ideas any more?'

Child 3: 'Wolf is trying to think of a plan.'

Child 4: 'Yes, but he can't do that because the cake has already been eaten.'

Transcript 14.: Example exchange of children when interacting with the IST.

This disagreement was solved by coding all utterances of this kind as PLS.

With the P×R scheme, we were able to pinpoint which perspective a user assumed and which frame he or she referred to. In contrast, as I argued in Section 3.4.2, both Sawyer's and Fine's annotation schemes only allow for limited annotation of communication in the context of interactive storytelling systems.

Sawyer's approach to in-frame and out-of-frame communication does not distinguish between player and person perspectives. With Sawyer's scheme, all utterances in our PLS, PLG, PES and PEG categories would have been annotated as level 3, as they are all out-of-character observations about story and game elements, which in Sawyer's scheme would amount to the play frame. Table 6 shows that in that case, 91% of all utterances would have been classified as level 3. This makes it clear that when using Sawyer's approach, we would not have been able to distinguish between the majority of utterances.

Fine's framework does not include blending between frames, which means that certain nuances would be lost when applying it to our communication data. This can be explicated by Transcripts 13 (coded as PLG) and 14 (coded as PLS). In Fine's scheme, the communication in these fragments would be classified as being in the context of the game and therefore as communication from a player perspective. Both fragments indeed show a player-perspective, but Transcript 13 refers to rules of the game, while Transcript 14 is about taking action in the fantasy world. In

Fine's classification, the latter fragment's reference to the story frame could not be made explicit.

Compared to the frameworks of Sawyer and Fine, the $P \times R$ scheme gives a more fine-grained view of the users' communication by giving insight into both the perspective of the user and the frame that is referred to. However, with respect to certain classification categories the scheme could be said to be overly fine-grained. Of the categories for CR, CG and PR communication, only one instance was found in our data: a child's tangible prop was insulted out of character by the other child, and she reacted with being 'angry' in character. This instance, labelled as CR, occurred in session 2 as only one of the 410 annotated utterances.⁶ Although theoretically possible, we found no instances of either the CG or the PR categories in our data. Note that CR and CG correspond to Sawyer's frame level 2; it is unknown how often this category occurred in his pretend play data (Sawyer, 2003).

There are also some distinctions the $P \times R$ scheme cannot make. We observed that references to the story made from a 'person' perspective (PES) included utterances that were explicitly addressed to one of the characters, for example 'Now, now Granny!' and 'Bad Wolf!' These were coded as PES, because they were neither spoken in character nor from a player perspective. However, such utterances did show great involvement in the story: the children addressed the characters as if they were 'real'. This distinguished these utterances from other PES utterances that showed a higher degree of detachment from the story world, such as comments on the events in the story ('Now it's getting scary.') and on the characters ('Wolf is not very smart.'), and cases where the children read aloud parts of the story text from the scrolls. The $P \times R$ scheme still seems not sufficiently fine-grained to capture these nuances. This, together with the finding that CR, CG and PR communication was lacking in the play sessions, encourages further research on our annotation scheme.

3.5 REFLECTIONS ON THIS CHAPTER

In this chapter, I addressed the first Research Question of this dissertation:

RQ1 What kind of meta-thinking do people do while co-creating stories?

To answer this question, I discussed two studies extending the work of Alofs et al. (2015), who created the Interactive Storyteller, an interactive version of the Virtual Storyteller. Using this system, children can co-create stories by adopting the roles of story characters and acting out stories together with computer-controlled characters. While children's understanding of narratives in picture sequences and TV shows has been researched (Shapiro & Hudson, 1991; Low & Durkin, 1998), few interactive storytelling systems have been investigated to understand children's interactions with such systems. The two studies I described were aimed at addressing this lack of insight by determining what the children think about their interaction while playing with the system.

⁶ As such, it made up only 0.24% of the total number of utterances in session 2 and was rounded to 0 in Table 9.

For the first study, we investigated how coherent the created stories were and which types of character actions contributed the most to story coherence, which was defined in terms of the causal connectedness of story elements. We found that the stories created by the children exhibited both global coherence, in the form of goal-action-outcome sequences (Trabasso et al., 1992) and local coherence in the form of immediate responses to other character's actions (Sawyer, 2003). However, the children's stories were only partially coherent. This can be explained by the improvisational nature of emergent narrative, where the interactive experience of the players is more important than the production of a coherent narrative. The children in this experiment spontaneously selected relatively many actions that had a strong potential for being causally connected, and used them to create goal plans. From this, we conclude that focusing on such actions in the creation of story worlds for emergent narrative may improve both the players' experience and the coherence of the created stories (cf. Swartjes and Theune (2009a)). The children switched between playing around in the storyworld and planning to advance the story. It seemed that the children took on different roles during their play sessions, but we could not deduce from just the story logs why this happened. To address this lack of information, we inspected the video recordings of the children's interactions to see how the children communicated about their interaction.

We performed a second study based on the recorded video material, and observed that the communication between the children, external to the system, plays an important role in establishing coherence of the created stories. This observation led us to investigate how children communicated about their co-creative interaction with the Interactive Storyteller. What users say during their interaction with a digital storytelling system can tell much about their meta-thinking, for example, how they relate to the characters and how engrossed they are in the storytelling activity. In the second study, we investigated the created stories and the utterances of the children during their interactions with the IST. We developed the $P \times R$ communication annotation scheme that combines ideas about framing from research on pretend play and role-playing games (Sawyer, 2003; Fine, 2002). This annotation scheme allows annotators to classify utterances and actions of users along the dimensions of *perspective* and *frame of reference*. For example, a child may propose (from a player perspective) an action for a character (the story frame of reference), or she may comment about the system's interface from a person perspective on the game frame of reference. We used the $P \times R$ annotation scheme to analyse the children's utterances and found that certain kinds of communication are prevalent among the children's interactions. At first, the children tended to discuss the workings of the system itself, yet, over time, their behaviour shifted to comments about the story from different perspectives. This can be attributed to the design of the IST: it does not particularly encourage in-character behaviour as it focuses on letting the children collaborate with the other player. Improvements in the user interface of the IST may further decrease the amount of comments on the system's mechanics and looks when the interaction would be made more streamlined. For example, the animations of certain actions could be made more attractive, as could the visualisation of the story world and interface elements. Players could then, instead, discuss the story

among themselves and make more elaborate plans. To further support the children going meta, the IST can be expanded with mechanics that allow more insight into possible consequences for actions. For example, the interface could show several possible storylines that might result from the player's choice. In turn, the player may then reconsider what she would rather see happen in the story and discuss this with her companion player.

We limited our analysis of interactive storytelling systems with the $P \times R$ scheme to the IST. Future work should look into the use of this scheme for the investigation of other similar systems. A comparison study between the IST and other systems could pinpoint deficiencies and improvements in either.

This chapter shows that children actively go meta when interacting with a storytelling system. They explain their actions to each other, to make their intentions clear and keep the story coherent, and give their opinion on what has happened and may happen. The meta-communication used by the children is valuable to their creative process. Not being limited to purely the character role enables them to express their feelings about the interaction and resolve possible conflicts. Furthermore, the children's going meta is indicative of parts of the interaction they might struggle with, such as shortcomings of the interface or the authoring of the story.

AGENTS THAT PERCEIVE AND ASSUME TO STEER STORIES

4

Where's the fun without complications?

Beckett
Vampire: The Masquerade – Bloodlines
Troika Games (2004)

In this chapter, I discuss how the meta matters to virtual agents of the Virtual Storyteller (vst).¹ This chapter treats Research Question 2 of this dissertation:

RQ2 How can virtual agents use meta-thinking to improve the generation of stories?

To answer this research question, I discuss techniques that virtual agents can use to overcome the limitations of the *strong autonomy* approach in interactive storytelling systems. This approach lets character agents develop a story's plot themselves by having them interact with each other, without a centralized manager as in the *strong story* approach (see Chapter 2). The Virtual Storyteller (vst) adopts the former, giving characters the freedom to follow their own plans. However, the way the vst was designed gave characters complete knowledge of the worlds they inhabit. This omniscience limits the stories the characters might create. For example, stories featuring deception are not possible, as characters are unable to hide from others, who will immediately know where they are. We created several expansions of the vst's capabilities to make it possible for scenarios featuring deception and incorrect knowledge to take place. The first of these is that, instead of having one complete picture of the world, the character agents have an additional knowledge base which comprises their *in-character* (ic) knowledge. Their omniscient world view is stored in a separate, *out-of-character* knowledge base. This allows characters to have a limited (ic) world view, enabling characters to perceive only those changes in the world that are visible to them. To ensure that the stories that result from this change in the perceptions of characters remain interesting, we have implemented the ability of characters to make assumptions about facts unknown to them. Moreover, the character agents can use their out-of-character knowledge to influence their in-character decisions. By going meta, the agents can steer the plot, for example by inciting conflict between their characters.

IN-
CHARACTER
(ic)
OUT-OF-
CHARACTER
(ooc)

In Section 4.1, I discuss storytelling systems similar to the vst, especially concerning their approaches to perceptions and character-based planning. I go

¹ This chapter is based on ten Brinke, Linssen and Theune (2014) and ten Brinke (2014).

into more detail about the VST's planning and its shortcomings in Section 4.2. In Sections 4.3 and 4.4, I discuss the mechanisms we designed to allow characters to perceive and assume. Section 4.5 contains a small evaluation of generated stories using these new implementations. In Section 4.6 I close this chapter by reflecting on the possibilities of the use of the newly designed perceptions and assumptions for characters.

4.1 PERCEIVING AND ASSUMING IN INTERACTIVE STORYTELLING

TALE-SPIN

Story generation systems that use a character-centric (strong autonomy) approach, like the VST, handle perceptions and planning in different ways. A classic example is TALE-SPIN (Meehan, 1977), one of the first story generation systems. Meehan already encountered the problem of perception as he overlooked the concept of 'noticing' when writing the inference rules for his program. This led to the 'mis-spun' tale below, in which Bill Bird, sitting on the river bank, did not see that his friend Henry Ant fell in the river and drowned. While this is exactly the opposite problem from that encountered in the VST, it highlights the necessity of having a perception system for characters.

Henry Ant was thirsty. He walked over to the river bank where his good friend Bill Bird was sitting. Henry slipped and fell in the river. He was unable to call for help. He drowned.

Transcript 15.: Example story generated by TALE-SPIN (Meehan, 1977).

OZ PROJECT

Another seminal work is the *Oz Project* (Bates, Loyall & Reilly, 1992). The Oz character agents run the typical three-step sense-think-act loop. The first step is to obtain raw sense data from the world. An integrated sense model keeps track of the agents' knowledge of the current state of the world. During the think step, the agent's next action is selected, based on its perceptions, current goals, emotional state, and other internal elements. The agents use a procedural reasoning system which has a library of plans. Every goal has one or more predefined plans that can be used to achieve it. All active plans and goals are stored in a structure called an active plan tree, which is updated after receiving sensory information. The perception system used by Oz is versatile and enables some of the possibilities we would like to have in the VST. Agents in Oz have an incomplete and possibly inaccurate view of the world, which makes it possible to deceive them by hiding information or by letting them misinterpret their perceptions.

FATIMA

FearNot! is an emergent drama system designed to teach children how to deal with bullying (see Section 2.1.3.1). Like the VST, its characters are autonomous agents that use partial order planning to select their actions. In *FAtiMA*, the agent architecture used by *FearNot!* and its successors (Aylett et al., 2006), the agents receive all perceptions, just like those in the VST. However, the stories all take place in single locations such as a class room or a school yard. In this context, the characters' omniscience does not lead to problems as it is believable for all characters to have the same perceptions.

Chang and Soo (2009) also outfitted character agents with social reasoning, enabling them to create an emergent narrative through their interactions. Characters can, based on their perceptions of others' behaviour, plan to influence the mental states of the others, for example, by stealing something they cherish or by disrupting their relationship with other characters. To ensure the emergence of interesting stories, Chang and Soo implemented a story facilitator module called the *Divine Intervention Operator Service* (DIOS), which adjust characters' personalities to the demands of the planner. For example, the DIOS could give a character extra information in order for her to hate another character, inciting a plan to let the first character plot against the second.

DIOS

In the *Actor Conference* system (Riedl & Young, 2003), character agents make use of partial-order planning. Via a blackboard architecture they contribute to an overall story plan, each making plans for its own character. Each agent has a limited plan space, but this is to reduce the complexity of the planning problem rather than due to limited perceptions. Plan spaces can be expanded if narrative generation is threatened by incompleteness.

ACTOR
CONFERENCE

The main question I address in this chapter is how agents can deal with the planning problem if their knowledge of the world is incomplete. None of the examined systems provide a solution for dealing with planning when characters do not, or rather, are not *supposed* to have all the information they need.

The problem of planning with incomplete information has been previously addressed outside the story generation domain, in particular in the context of planning for robots. Several solutions involve the use of default assumptions (Weser, Off & Jianwei Zhang, 2010; Davis-Mendelow, Baier & McIlraith, 2012). Such work is generally aimed at maximizing the plausibility of assumptions and minimizing the costs of plan failure. However, occasional plan failure is not a big problem in a story generation context. Characters in stories frequently run into situations that do not turn out as planned. How they deal with the problem is often an important part of the story. A key question therefore is how to make the assumptions believable, and how these assumptions can be used to steer the story in an interesting direction. In the interactive storytelling system of Cavazza, Charles and Mead (2002, see Chapter 2), the characters use a Hierarchical Task Network (HTN) planner that defines how goals are decomposed into sub-goals and possible plans to achieve them. The leaves of these plan trees are terminal actions with preconditions on their use. The preconditions for a terminal action may not hold, due to the activities of another character or user interference, causing action failure. When a plan fails, for example due to an object being stolen by the user, the agent will find the next applicable plan in its HTN and try that instead.

4.2 LIMITATIONS OF OMNISCIENCE FOR STORY GENERATION

The stories in the VST emerge from the actions of autonomous character agents in a simulated story world. The character agents of the VST have a dual role as actor and character (Swartjes & Theune, 2009b). As characters, their main responsibility is to behave in a believable fashion. As actors, the agents bear a responsibility for creating an interesting plot. The character agents go through a deliberation cycle

in which perceptions of actions and events lead to beliefs. These beliefs then enable new actions or lead to the adoption of goals for which actions can be planned. When a character performs an action, the story world is updated and all agents are informed of the change; they ‘perceive’ it.

At the core of the character agents’ deliberation cycle is a *partial-order planner* (POP) that can determine which actions to perform to reach a particular goal. The POP used in the VST tries to make a complete plan from start to finish, which will fail if crucial knowledge is missing. Due to the closed world assumption used in the VST, any fact not explicitly known to be true is assumed to be false. If the location of an object is unknown, the planner concludes that making a plan to get that object is impossible.

Multiple story domains have been defined for the VST, including a classical fairy tale domain with knights and princesses, a story world inspired by Little Red Riding Hood, and a pirate domain (see Section 2.2). For the current research, we made a variation of the pirate world PLUNDER. The setting is a pirate ship with various locations; the main characters are O’Malley, the ship’s cat, and Scurvy, the bilge rat. Scurvy wants to eat the cheese that is located somewhere on the ship and O’Malley has the goal to catch Scurvy.

While the virtual agents were already able to autonomously create different stories (see Section 2.2.1), their omniscience of the story world limited the possibilities for story generation. For example, nothing could be hid from character agents, as they always had complete knowledge about the state of the world. We addressed this shortcoming by re-designing the characters’ perceptions (Section 4.3) and enabling the characters to make assumptions about the story world (Section 4.4).

4.3 PERCEPTIONS IN THE VST

As explained above, omniscient characters make it impossible to generate stories with suspense and deceit. To overcome the disadvantages of omniscience in characters, we created a new perception system based on the distinction between character (*in-character*, IC) and actor (*out-of-character*, OOC) knowledge. Here we take our inspiration from Gordon and Lent (2002), who present a discussion on the use of virtual humans as participants (characters, IC) versus virtual humans as actors (OOC) in virtual environments. The choice between the two affects how agents perceive the world: virtual participants should only perceive those parts of the world they can realistically perceive, while actors should be omniscient. However, an actor should still be aware of the realistic bounds of his perception to remain believable. Plans characters make can then be based on IC knowledge, whereas the actor’s OOC decisions will be based on the full and accurate knowledge they require. We split the knowledge base for characters into parts for the character component (IC) and the actor component (OOC).²

² For efficiency, the VST only stores facts that differ between IC and OOC knowledge in different knowledge bases; the information that is the same in both views is stored in one knowledge base which is shared by the views.

In this section, I describe the set of perception rules characters use to perceive their world and how their IC knowledge is updated. First, I explain the rules used for visibility in Section 4.3.1; other senses are left for future work. These rules have to be applied to incoming perceptions to decide which ones to add to the IC knowledge base of the character performing the perceptions. Once the characters are keeping proper track of what they can see and know, they can be made to use only that knowledge for IC plans. Section 4.3.2 deals with the re-design of the characters handling perceptions. In Section 4.3.3, I explain how the characters' initial knowledge of the world is set up.

4.3.1 *Visibility rules*

The rules for visibility are as follows. Characters, objects and actions at the same location as the character can be seen by the character. By making use of path relations between locations, we can determine adjacent locations of which the character might reasonably have a view. The character has a view of other locations if there is a direct path to that location that is not blocked by a closed door. Objects in a container can be seen from the location of the container if it is open. These rules are quite basic, but they can easily be expanded.

For perception of an object, the rules work in three steps. The first step checks whether an object is at a location, either directly or indirectly, for example, if it is being held by someone. The second step checks from which locations the object is visible. The final step checks if an object is visible for a character, using the previous two steps and knowledge about the character's current location.

Visibility rules are an important part of the story world description. The responsibility of aligning them with the story world is left to the author of the story domain. On one hand, this is a downside as it requires more work from the author. On the other hand, it does allow for more flexibility. For example, if an author wants to add magical crystals that allow a character to see distant worlds, she can. It would not be possible to simply capture such creative solutions in a general system that works for all story worlds.

4.3.2 *Dealing with perceptions*

When events take place in the storyworld, the state of the world is updated. The first thing a character agent should do with a world update is handle it ooc so the actor's knowledge stays up to date. The perception is then analysed using the visibility rules to see whether the character gets to see it IC. An action is visible for a character if the character performing that action is visible. If this test returns that the object, character, or action is not visible, the perception should not be added as a belief to the character's IC knowledge base. Because the characters only get updates (new perceptions) when changes occur in the storyworld, we do not discard the facts that are not perceived IC. Instead, we store them ooc, so that the character can add them if they become available at a later time. If we keep track of things the character missed, it becomes much easier to re-test visibility later in the story. For example, a character hiding behind a desk may easily be

Figure 15.: Schematic map of the ship in the story world.

overlooked by another character from the front side of the desk, yet if he moves to the back, he will spot the other character. The result of this is a collection of facts that either a character believes but are not true or things that have occurred but the character has not seen yet. Whenever a new perception is received, previously missed facts are re-tested for visibility. This amounts to taking care of two different cases. The first case concerns beliefs that must be added due to the effects of the new perception. For instance, if O'Malley sees that Scurvy is at the deck, this means he will also conclude that Scurvy is no longer in the galley. The second case involves facts that have to be removed from the knowledge bases because they are no longer true. For example, Scurvy believes *ic* that the cheese is in the galley. While Scurvy may believe *ic* that the cheese is there, the visibility check relies on the *ooc* knowledge base, as it holds the objective world state which overrules *ic* beliefs. This results in dropping the false belief of the cheese being in the galley, as it is not there according to Scurvy's *ooc* knowledge.

4.3.3 Initial character knowledge

In the *vst*, characters get their initial knowledge of the world at the start of a run of story generation. Without this, they will have an empty *ic* knowledge base and will be unable to plan. We give each character a separate copy of the actual state of the world, but allow this to contain false information. This grants flexibility to authors of characters, as they can change a few things in the characters' initial knowledge and use this to manipulate their behaviour to suit the story. For example, we can make sure O'Malley does not know *ic* where Scurvy is at the start of the story, so that he cannot plan to catch Scurvy until he first sees him. At the same time, it allows us to give Scurvy the false belief that the cheese is in the galley.

To test the visibility rules, we created a scenario in which Scurvy has a goal to get the cheese and O'Malley wishes to catch Scurvy. Figure 15 shows the schematic layout of a pirate ship which is the basis for the story world. Scurvy was given the initial belief that the cheese was in the captain's cabin, leading him to go to this location. After opening the door to the captain's cabin, he noticed the cheese

was not there. As expected, Scurvy did not know where to look for the cheese and could not make a plan to get it. Without any other goals to pursue, he did nothing and got caught by O'Malley. In the next section, I explain how we let characters use assumptions about the story world to overcome this problem of planning with missing knowledge by looking ahead.

4.4 ASSUMPTIONS IN THE VST

Limiting the character to using only IC beliefs for planning introduces a new problem. If the character has the goal to acquire an object, he is unable to make a plan to do so if he does not know where the object is. As the POP requires this knowledge to build a plan, planning will not be possible, leading the character to drop the goal. This of course does not lead to interesting stories. What is necessary is a way for characters to fill in the missing information.

Multiple approaches could be used to give the characters more information. Implementing additional senses like hearing or smell would provide more ways for the character to receive a perception. What we required is for the characters to be able to make plans based on incomplete knowledge, without making drastic changes to the planner or the rest of the system. To this end, we created assumptions operators, which I discuss in Section 4.4.1. In Section 4.4.2, I describe the necessity for preconditions of assumptions. Sections 4.4.3 and 4.4.4 show how assumptions can be executed by characters and how they can be used by characters to control the plot, respectively.

4.4.1 Assumption operators

As a result of their new abilities for perception, characters can now have incomplete or inaccurate knowledge of the world. This entails that the closed world assumption no longer holds for their IC knowledge.³ When a character does not know the location of an object, it should assume that the object is somewhere that it cannot currently see, instead of assuming that the object is not located anywhere. To make this possible, we implemented an assumption operator that works similarly to other planning operators. The assumption operator does not need to provide the correct location for the object. If it did, the situation would not be much better than having omniscient characters. Instead, it should be possible for a character to search in the wrong location, possibly leading to dramatic situations. The planner has no trouble using false information to make a plan. At some point during the execution of the plan, a perception will trigger and the false information will be removed from the knowledge base. At this point the plan will become invalid and a new plan will be made if possible.

Having separate assumptions for every object that a character can lose track of would soon become unmanageable for larger story worlds, so a general operator is needed. A schema for a general `AssumeObjectAtLocation` assumption would have three arguments: *agens*, *patiens* and *location*. The *agens* is the character making

³ However, the closed world assumption does still hold for the ooc part of the agents.

the assumption. The patients is the object the assumption is about, and the location is where the object is assumed to be. The effect of the assumption operator is a fact that puts the patients at the location, but only in the character's mind (ic).

4.4.2 *Preconditions*

If we allowed the planner to make assumptions without proper preconditions, it would simply assume any object is at the character's location because this would provide the shortest plan when trying to find it. Moreover, if the planner is allowed to make assumptions about an object without restrictions, it will assume the object is at a more convenient location even if the actual location is known but going there would result in a longer plan. Therefore, a couple of rules are needed to make assumptions work in a believable manner. First of all, if a character knows the location of an object, it should not be allowed to make any assumptions that contradict this knowledge. Second, the object must exist; in other words, the actor must have ooc knowledge about the location of the object. It is a trade-off where we have to give the character some information it couldn't logically know yet, because we do not yet have a method for the character to determine that the object exists. Without this constraint the character might begin a search that will never end. If the object does not exist, it may be possible to add the object to the world through improvisation by using a framing operator, as described by Swartjes and Theune (2009b, see Section 2.2.1). Finally, the object cannot be at the character's current location. If it was, the character would have seen it.

Even given these preconditions, the character will end up walking back and forth between two locations unless he runs into the object he is looking for by chance. To make sure the character searches in a more believable manner, we take our inspiration from the Oz system, in which agents use current sensory information as well as past knowledge to select possible plans (Bates et al., 1992). Lyotard, the cat in the Oz simulation, has multiple strategies to find food. He can try the most likely location, his bowl, he can meow for attention or he can try a location where he remembers seeing food. The characters in the VST also keep a memory of past events by storing beliefs and perceptions in episodic memory. If we make use of these memories, we can find out where a character has been and how long ago this was. Using this information, we can exclude a location as a believable search location if the character has been there too recently. In our implementation, we let characters only search in a location when it has never been visited or when it has been visited more than 10 turns ago in the simulation.

4.4.3 *Executing assumptions*

Using the new rules, we implemented the `AssumeObjectAtLocation` operator by implementing assumptions as proper operators that have to be executed. Unlike other operators, they do not affect the actual world; they only change the internal world representation of the character that makes them, so their effects are ic only. The character stores the effects of the operator directly as beliefs. Whenever a character executes an assumption operator there is a chance that the resulting

belief does not correspond to the actual world state. However, the mechanism described in Section 4.3.2 allows the characters to deal with such false beliefs. This is achieved by informing the interpretation module of any false assumptions, so it can trigger the appropriate perception when needed.

With these changes, the character agent properly executes assumptions and also retracts them if they prove to be false. Now, instead of giving up when unaware of an object's location, the character will search for the object methodically, starting with the nearest possible location within the constraints from Section 4.4.2, because this will provide the shortest plan. Eventually the character will find the desired object and can continue with the rest of his plan.

Whether or not the resulting story is interesting is left to chance. During his search, the character might encounter other things he was unaware of. He may run into a conflict he was not expecting. But then again, he may also get lucky and guess right on the first try, avoiding any trouble.

In small story domains we can manipulate the location of characters and objects to ensure that something interesting will happen. But in bigger and more complex scenarios, controlling the location of objects and especially characters that way will not be possible. To circumvent this problem, we can take assumptions one step further.

4.4.4 Plot control

When a character makes an assumption, this is an ideal moment to take some control of the plot without impacting the character's believability. The actors in the VST should not try to reach their character's goal, but to create an interesting story. A story world may have opportunities for events that, while interesting for the audience, are not in the interest of any of the characters' goals. If the agent has OOC knowledge about these opportunities, it can direct the character towards them, instead of directly towards the character's goal. At a moment when a character no longer knows where to go, the actor can steer the character towards conflict without limiting the character's autonomy. To maintain believability, the actor will have to keep in mind what the character does know.

In short, by making use of the actor's OOC knowledge, we can manipulate the character's choice and send it in a direction that will put it in an interesting situation. The underlying idea is that an author usually does not send a character to a location without a reason. The character may find what it is after, or it may find something else, positive or negative. Eventually, the actors need some sort of meta-goals to help make the story more interesting, but for now we use a simpler, more direct approach to give an impression of how this could work. By making separate assumption schemas for each phase of the story, we can add preconditions to either introduce conflict or resolve it.⁴ In the first half of the story we can alter the assumption so it guides the character towards a conflict. After the main conflict has been resolved we can use a second variation of the assumption to tie up loose ends and finish the story.

⁴ This is similar to making actions available to a character depending on the story phase, as proposed by Tommassen (2009).

4.5 EVALUATION

In this section, I discuss how the characters' new capabilities of perception and assumption can be used to author interesting stories. Like before, we use a story world in which Scurvy wants to obtain the cheese. When the story starts, he has no knowledge of the location of the cheese, so he will have to search for it. To make the stories more exciting, we want to make sure that Scurvy runs into O'Malley, who will adopt the goal to catch Scurvy as soon as he can see him.

4.5.1 *Expanding the story world*

We use an expanded version of the story world from Section 4.3.3. Previously, Scurvy only had the goal to get the cheese and therefore he would just ignore O'Malley's presence. To prevent this, we add an escape goal, which will make Scurvy flee once he sees O'Malley. Once the escape goal has been triggered, Scurvy will try to run from O'Malley, but since they both move at the same speed, Scurvy will either get backed into a corner or he will find the location of the cheese. In both cases, Scurvy will need a way to keep O'Malley busy, otherwise he will not have enough time to get away or pick up the cheese. For this purpose we add a *distract* goal, which can be achieved by either of two new actions reminiscent of the Tom & Jerry cartoons. The rat can drop a cannonball on O'Malley's foot, preventing him from following, and he can kick over a bucket of soapy water, which will make the deck slippery. For O'Malley there are two new actions that he performs in response to Scurvy's distracting actions: dancing around on one foot while screaming in pain, and cleaning the deck. We also give O'Malley the goal to swab a deck if it is slippery, to keep the captain from getting angry with him because of an untidy ship. As a consequence, if Scurvy makes the deck slippery, O'Malley is forced to swab the deck. The captain is not in the story, he is just the motivation for this goal. Scurvy also gets some new goals that will help him avoid being captured. Finally, Scurvy gets a new action that lets him use the tunnel to get to his lair.

To make the story progress the way we want it to (as discussed in Section 4.4.4), we add two assumption operators. The first makes Scurvy assume the cheese is at the location where O'Malley is, by using the actor's ooc knowledge. This can be used early in the story to ensure that Scurvy will run into O'Malley and begin the conflict. The second assumption lets Scurvy assume the cheese is at the location where it actually is. This can be used at any point after O'Malley has been distracted at least once and thus the conflict has already taken place. This guarantees that, from then on, the story will start to work towards a resolution.

4.5.2 *Example stories*

In our first story, we let Scurvy start at his lair and O'Malley at the galley. The cheese is placed in the captain's cabin. Scurvy immediately takes on the goal to find a piece of cheese. Transcript 16 shows how the story plays out. Lines describing characters' perceptions and assumptions are shown in italics. The story is not very fluent, as we did not use the language generation component of the VST for narration.

Scurvy wants to get the piece of cheese.
Scurvy assumes the cheese is in the galley.
 Scurvy walks to the cargo hold via the door.
 Scurvy walks to the gun deck via the ladder.
 Scurvy opens the door to the galley.
Scurvy can see the cheese is not in the galley.
O'Malley sees Scurvy.
 O'Malley wants to catch Scurvy.
 O'Malley walks to the gun deck via the door.
 Scurvy wants to avoid O'Malley.
 Scurvy walks to the crew's quarters via the door.
 Scurvy wants to distract O'Malley.
 Scurvy walks to the gun deck via the door.
 O'Malley walks to the crew's quarters via the door.
 O'Malley walks to the gun deck via the door.
 Scurvy picks up the cannonball from the gun deck.
 Scurvy drops the cannonball on O'Malley's foot.
Scurvy assumes the cheese is in the captain's cabin.
 Scurvy walks to the deck via the ladder.
 O'Malley dances around on one foot, screaming loudly.
 Scurvy opens the door to the captain's cabin.
 Scurvy walks to the captain's cabin via the door.
 Scurvy picks up the piece of cheese from the captain's cabin.
 O'Malley walks to the deck via the ladder.
 Scurvy wants to hide with the piece of cheese at Scurvy's lair.
 Scurvy walks to the deck via the door.
 O'Malley walks to the captain's cabin via the door.
 O'Malley walks to the deck via the door.
 Scurvy walks to the gun deck via the ladder.
 Scurvy walks to the cargo hold via the ladder.
 O'Malley walks to the gun deck via the ladder.
 Scurvy walks to the brig via the door.
 O'Malley walks to the cargo hold via the ladder.
 Scurvy crawls to Scurvy's lair via the small hole.
 O'Malley walks to the brig via the door.

Transcript 16.: An example story generated by the VST with the use of the
 new perception and assumption capabilities of the agents
 (shown in *italics*).

Since Scurvy does not know where the cheese is, he has to make an assumption
 about its location. As the story has just started and no conflict has taken place yet,
 the agent uses its ooc knowledge about O'Malley's location (the galley) to assume
 that the cheese is at that location. As soon as Scurvy opens the door, O'Malley
 discovers him and adopts the goal to catch him. Scurvy tries to run away, but soon
 finds himself cornered in the Crew's quarters. He then decides he has to distract
 O'Malley in order to get away and uses the cannonball to achieve this. As O'Malley

is now distracted, Scurvy can get back to the goal of finding the cheese. Scurvy has found out that the cheese is not at the galley, so he makes a new assumption. Since the conflict has now taken place, a new story phase is entered in which the agent's ooc knowledge can be used to guide Scurvy towards the cheese. This helps to maintain the pace of the story instead of dragging it on with uninteresting searches in wrong locations. Scurvy now goes straight to the cheese's actual location, the captain's cabin. He grabs the cheese and can then adopt his final goal of taking it back to his lair. In the meantime, O'Malley has recovered from his sore foot, but as Scurvy is constantly moving, O'Malley is unable to catch him before Scurvy makes it safely back to his lair. There O'Malley can no longer reach him, so the story ends.

Our second example story takes a different course (see Transcript 17) because we change the starting locations of the characters: Scurvy begins on the poop deck and O'Malley begins on the gun deck. The cheese is still in the captain's cabin and, once again, Scurvy starts off by making a false assumption. O'Malley starts chasing him and Scurvy is forced to retreat to the poop deck. Unable to move further in this direction, he has to think of a distraction again. This time he spots the bucket of soapy water on the deck and decides to kick it over. O'Malley is obligated to clean the deck, giving Scurvy a chance to go for the cheese. Again, Scurvy uses ooc knowledge to assume where the cheese is, because the conflict has already taken place. O'Malley finishes his task just as Scurvy is on his way back from the captain's cabin with the cheese. He chases Scurvy again, but, as before, the quick little rat escapes through his tiny tunnel out of O'Malley's reach.

In the second story, Scurvy's first (false) assumption had little impact on the story because he had no choice but to head to the deck. But as the bucket of soapy water was now located much more conveniently than the cannonball, he used that in his plan to distract the cat.

The generated stories illustrate how the split between ic and ooc knowledge can be used. The characters will always try to make use of their ic knowledge and base their plans on that so long as the information is sufficient to make a complete plan. Whenever this is no longer the case because a belief proved to be false, the actor knowledge is used in combination with past beliefs from the episodic memory. The latter helps the characters to make believable choices as they will not search in places they have recently visited. The former is used to direct the characters towards interesting places depending on the current state of the plot. The stories also show that characters can be deceived, as in fact the actor component deceives the character and gives it false beliefs.

One problem with the implementation of the assumption and perception operators is that they are not narrated in the vst. This makes it difficult to see when and why a character makes an assumption and starts to behave differently. Eventually there should be a mechanism to determine which perceptions and assumptions are important to narrate. Any perception or assumptions that reveals a fact that is a precondition for a goal would be a likely candidate to be narrated as would any perception or assumption that contradicts a character's beliefs.

Scurvy wants to get the piece of cheese.
Scurvy assumes the cheese is on the gun deck.
 Scurvy walks to the deck via the ladder.
Scurvy can see the cheese is not on the gun deck.
 Scurvy wants to avoid O'Malley.
 Scurvy walks to the poop deck via the ladder.
O'Malley sees Scurvy.
 O'Malley wants to catch Scurvy.
 O'Malley walks to the deck via the ladder.
 Scurvy wants to distract O'Malley.
 Scurvy walks to the deck via the ladder.
 O'Malley walks to the poop deck via the ladder.
 O'Malley walks to the deck via the ladder.
 Scurvy kicks over the bucket of soapy water.
 O'Malley is distracted by the slippery deck.
 O'Malley wants to clean the deck.
Scurvy assumes the cheese is in the captain's cabin.
 Scurvy opens the door to the captain's cabin.
 O'Malley walks to the poop deck via the ladder.
 O'Malley picks up the swab from the poop deck.
 Scurvy walks to the captain's cabin via the door.
 Scurvy picks up the piece of cheese from the captain's cabin.
 O'Malley walks to the deck via the ladder.
 Scurvy wants to hide with the piece of cheese at Scurvy's lair.
 Scurvy walks to the deck via the door.
 O'Malley swabs the deck.
 Scurvy walks to the gun deck via the ladder.
 Scurvy walks to the cargo hold via the ladder.
 O'Malley walks to the gun deck via the ladder.
 Scurvy walks to the brig via the door.
 O'Malley walks to the cargo hold via the ladder.
 Scurvy crawls to Scurvy's lair via the small hole.

Transcript 17.: Another example story showing the new perception and assumption capabilities of the agents (indicated in *italics*).

4.6 REFLECTIONS ON THIS CHAPTER

In this chapter, I addressed the second Research Question of this dissertation:

RQ2 *How can virtual agents use meta-thinking to improve the generation of stories?*

To answer this question, I discussed how we improved the Virtual Storyteller by turning omniscient characters into more believable ones whose knowledge of the story world is limited by what they can realistically perceive. The purpose of this endeavour was to enable the generation of previously impossible stories and

to make the generated stories more interesting. The character agents in the VST received separate knowledge bases for in-character (IC) and out-of-character (OOC) knowledge. The first stores the character's current knowledge of the story world and the second stores the actual world state for the actor. The character agents decide for themselves which perceptions to adopt as IC beliefs. To this end, a set of visibility rules was implemented. In the current version only the sense of sight was implemented. The rule system can be expanded as the story world requires and other senses can be added to it at a later time. The visibility rules limit the characters' knowledge about the story world, so we had to provide a means of filling in missing information. To do this an assumption operator was implemented that allows a character to make a believable guess about the missing facts and use it in planning. If the assumption is wrong, the character simply tries again until it gets it right.

Since random searching does not add much excitement to the story, we implemented plot control options through the assumption operators. When the character has to decide where to go, we have a good opportunity to direct the character towards a potentially interesting situation. Making use of OOC knowledge, the character can be directed either to the correct location or to a location that adds some conflict and excitement to the story. Thus, the authors of the storyworld can provide the agents with the means to improve the generation of stories. This effort in co-creating a story relies on the author going meta, reflecting on opportunities for agents to deceive each other. It is up to the agents to then go meta, under the circumstances specified by the author, and steer the story in a certain direction.

In the example stories we used assumptions to direct the hero, Scurvy, directly towards the villain, O'Malley, in order to enforce conflict between the two. After the conflict happened, another assumption guided Scurvy towards his goal in order to resolve the story. These generated stories show that the perception system and the assumption operators work as intended. The characters now act on beliefs that they could logically have instead of knowing and reacting to every change in the world as it happens. The guided assumptions let the characters make wrong decisions and can be used to force conflicts to happen, even if a character would have avoided them if he had known everything.

This work on perceptions and assumptions of characters in the VST is far from complete. There are four main facets of the new mechanics for perceptions and assumptions that can be investigated. Firstly, the assumption operators used in the example stories were tailored specifically to the story domain. The only interesting locations were the location of O'Malley or the actual location of the cheese. In the future, more options should be added. For this, we would like to use a more high level approach. Our approach would benefit from a higher level perspective. By giving the actors meta-goals to achieve, they can decide which are the most interesting assumptions, similar to the double appraisal mechanism in FearNot! (Aylett et al., 2008). For example, if the meta-goal is to start a conflict, the actor can pick from a number of available conflict descriptions and try to fill in the assumption in such a way that the character's resulting plan will achieve the preconditions that will cause the conflict. An example conflict may occur when

two characters are searching for the same treasure and one already has a piece of the map. Then, we can use the assumption to guide the other character to a second piece of the map which will give him the goal to find the treasure as well. This would require further adaptations to the planner so it can work with multiple goals on IC and OOC levels simultaneously. Actors would also have to keep track of the goals of other characters so they can anticipate their moves. This requires developing a theory of mind for the character agents (Chang & Soo, 2009; Sindlar, Dastani, Dignum & Meyer, 2010).

Secondly, we have only implemented the sense of sight. Implementing at least the other classic senses (hearing, touch, smell and taste) would increase the number of stories that can be told. Hearing in particular will become more important when the characters are given the ability to speak and exchange information. Using this, we can also create scenarios that involve eavesdropping, which is commonly used in stories.

Thirdly, the decision on which perceptions a character will receive is currently purely on rule-based. Sometimes, things may be visible according to the perception rules, but it may be more interesting for the story if the character does not notice them. For example, a character who has his pockets picked will currently always notice this according to the rules, but for the story it might be better if it is only discovered later. For these cases it would be good if there was a method to determine how obvious an effect is to a character, based on factors like the size of an object, the loudness of a sound and the character's focus of attention.

Fourthly, our evaluation of the generated stories was limited to manually inspecting them. Subjective comparisons between stories which were generated either with or without the new perceptions and assumptions could provide valuable insights into how perceptions and assumptions may work best. While multiple evaluation procedures exist for such comparisons, such as those by Riedl and Young (2005), Peinado and Gervás (2006), and Roth and Koenitz (2016), comparing such stories is a difficult task, as they may take wildly different turns depending on the choices of the characters.

The work I discussed in this chapter revolved around the behaviour of virtual agents for story generation. To elevate this work to the realm of interactive storytelling, the most challenging hurdle that needs to be overcome is the modelling of users. Whereas character agents have out-of-character knowledge about their peers, they cannot tap directly into users' minds. As explained above, this calls for a theory of mind which characters can use to reason about users' intentions. For example, in order to keep things secret from users, character agents require knowledge about the users' thoughts. Physically hiding from users may be implemented straightforwardly, given sufficiently rich rules for perception; characters would then reason to prevent being perceived by users. Obfuscating or deceiving users, for example, by telling them lies, is a more challenging task (Ryan, Summerville, Mateas & Wardrip-fruin, 2015). Implementing this would require the agents to actively keep track of the things they believe users to know; even more so when they try to scheme against users, which involves long-term plotting. As I briefly discussed, dialogues between characters in the VST are still fairly limited and straightforward. Dialogue models should first be implemented in the

VST before the agents will be able to deceive users. Nonetheless, the approach I presented here offers many possibilities to improve interactive storytelling systems. The character agents could use their out-of-character knowledge about users in the same way they currently do with their fellow agents. For example, the character agents can take into account the expectations of users about the outcome of a story, ensuring it has a happy or sad ending. Alternatively, the character agents might choose to do the inverse of what users suspect or expect, increasing the suspense and surprise of stories. By going meta, the agents may be able to reason about the story and, with further work, also about the users, improving the way interactive stories play out.

PART III

META MATTERS IN SERIOUS GAMING

In this part, I look at how the meta matters in the field of serious games, from the perspective of both human and virtual agents. This part follows the same structure as Part II.

In Chapter 5, I provide both an overview of existing serious games for social skills training and an explanation of LOITER, the game we developed for the studies in the following chapters.

In Chapter 6, I answer Research Question 3:

How can serious games stimulate meta-thinking to increase their effectiveness?

I present meta-techniques that present information to players of LOITER, a serious game for social skills training. These techniques are aimed at giving insight into the thoughts of virtual characters, letting players see and reflect on the effects their actions had on the characters.

In Chapter 7, I answer Research Question 4:

How can virtual agents use meta-thinking to adapt serious games to the user's competence?

I discuss how virtual agents can use meta-information about players' competence of LOITER to adapt the game's difficulty. When players do well, the agents change their behaviour to be less cooperative, making it more difficult for players to make them collaborate. To keep their behaviour coherent, the agents use meta-techniques to motivate their change in attitude.

What value a tale if it is not instructive?

Monok Ochem
House of Chains
 Erikson (2009)

THIS chapter provides background material on serious games for social skills training for Chapters 6 and 7.¹ It comprises two parts. In Section 5.1, I discuss existing serious games for social skills training. In Section 5.2, I discuss the serious game we developed and studied for this part of the thesis: LOITER, LOitering Teenagers, an Emergent Roleplay.

5.1 SERIOUS GAMES FOR SOCIAL SKILLS TRAINING

Chapter 1 provided a short introduction to the field of serious games, discussing their intentions and implementations. In this section, I discuss existing serious games related to my research on improving social skills.

Attaining and maintaining good social skills is of importance to many professions (D. W. Johnson & Johnson, 1990; Meier, DiPerna & Oster, 2006; Morgeson, Reider & Campion, 2005). Since the early 1970s, a lot of serious games and simulations have been developed to let people experience and try to deal with difficult social situations (Duke & Kriz, 2014). These often take the form of role plays without much technology involved. As explained above, I focus on digital serious games for social skills training. In their taxonomy of serious games devoted to personal and social learning, and ethics, Pereira et al. (2012) divide this broad field of research into three main branches of serious games. The first of these contains games for teaching soft skills, both personal and interpersonal skills, for example, adapting to changes and managing conflicts, respectively. The second branch focuses on ethics, while the third branch contains games meant to improve social awareness of people, for example, about their social responsibilities. With this distinction, Pereira et al. sketch a good overview of the different kinds of serious games for social training. However, they do not provide an in-depth discussion of serious games meant to train interpersonal skills. In this section, I focus on those systems which feature virtual agents with which people can interact in different social settings, like LOITER, our serious game. Some of these systems were already discussed in Chapter 2, as they also build on research on interactive storytelling.

¹ This chapter is based on Linssen (2012) and Linssen, de Groot, Theune and Heylen (2015).

5.1.1 Tactical language training systems

TACTICAL
LANGUAGE
TRAINING
SYSTEMS

One prominent line of systems meant to simulate interpersonal interactions with virtual agents is that of the so-called *tactical language training systems* (TLTS) developed at the University of Southern California, which are meant to train negotiation skills of military personnel (Swartout, 2010). These systems feature a virtual reality setup in which the player is required to negotiate with embodied virtual agents representing, for example, team members of his squad, civilians of other cultures and enemies. The user's goal in these negotiation systems is to bring existing conflicts to a peaceful resolution through investigating the case at hand, understanding which stakeholders are involved, and finally negotiating a solution that works for the involved parties. The pedagogical goals of the TLTSs can be authored in three different ways: including them in the virtual world's dynamics, integrating them in the intentions of virtual characters, or by letting a drama manager decide which events should occur to support those pedagogical goals. The first can be brought about by making the pedagogical goals congruous with the explicit game goals. For example, in a certain scenario, players are required to establish a virtual character's trust to progress through the game and by making it clear to the player that failure to do so will have its consequences. In other words, it can be done by designing the game so that the player is forced to go about fulfilling his goals in a particular manner. For the second way of authoring pedagogical goals, authors of a scenario for one of the TLTSs can assign intentions to virtual characters with the goal of letting players learn about different personalities and social customs. Through reasoning about the player's mental model the character can then infer how to act. Finally, a drama manager can watch over the proceedings during interaction between users and a TLTS and act on them when necessary, as in *Mimesis* and *IN-TALE*, which strive to guide the player towards situations from which she can learn.

5

MISSION
REHEARSAL
EXERCISE

The first of the TLTSs developed at the University of Southern California was the *Mission Rehearsal Exercise* (MRE) project (Rickel et al., 2002). In this project, a demonstrator was created in which the player assumes the role of a lieutenant in an international peacekeeping scenario. In it, he can interact with three agents: a sergeant and a medic on his team as well as with a local woman, see Figure 16a. The player's squad has been in a car accident with the woman and her son and it is up to the player to decide how to handle this situation. His squad needs to move ahead, but the woman needs to be comforted, because her son needs help after incurring injuries due to the accident. Thus, the player faces some choices on how to direct his team members which influence the outcome of the story. For instance, if he does not choose to offer any help to the woman and her son, the woman will start complaining and weeping. Alternatively, if he does offer help, his accompanying team may complain. Through negotiation with the virtual characters, the player needs to find an approach that solves this conflict situation. The three characters in MRE use cognitive models of emotion to govern their behaviour. They do so through a loop consisting of appraising the environment, goals and beliefs, after which coping behaviour directs how to act on those (Gratch & Marsella, 2001).

Figure 16.: Four tactical language training systems developed at the Institute for Creative Technology (adapted from Rickel et al. (2002), Traum, Swartout, Marsella and Gratch (2005, 2008), W. Johnson and Friedland (2010), respectively.)

As a follow-up to the MRE project, the Stability and Support Operations – Simulation and Training (SASO-ST) project was established (Traum, Swartout, Marsella & Gratch, 2005). SASO-ST focused explicitly on negotiating with a single virtual agent in a TLTS, see Figure 16b. This project integrated *Thespian*, an architecture designed to support the encoding of pedagogical goals (Si, Marsella & Pynadath, 2005). It relies on *PsychSim* as a framework for the social behaviour of its agents (Marsella, Pynadath & Read, 2004). *PsychSim* outfits agents with mental models that give them the capacity to reason about each other's behaviours and goals as well as about the effects of their behaviour on their relationship with other agents and with the player. This way, they can produce more natural, social reactions and plans. Traum et al. extended the *Thespian* architecture with four enhancements: appraisal of the negotiation; trust modelling; negotiation strategies; and behaviour explanation.

Moving on to more complex scenarios, SASO-EN (Stability and Support Operations – Extended Negotiations, see Traum et al., 2008) incorporated an extra agent into the simulation, thus forcing the player to bring a multilateral negotiation to a satisfactory outcome, see Figure 16c. This enabled more negotiation tactics, such as convincing one agent to cooperate, which would then help the player to convince the remaining agent.

SASO-ST

THESPIAN

PSYCHSIM

SASO-EN

OLCTS Concurrent with the creation of and research performed in the SASO systems, another TLTS called the *Operational Language and Culture Training System* (OLCTS) was developed by W. Johnson and Friedland (2010), see Figure 16d. This system was originally built on the PsychSim framework but discarded it in favour of one that was more manageable in terms of agent authoring. It adopts roughly the same approach as that of the SASO systems, but as it is a commercial product, not much is published about its inner workings.

BILAT *BiLAT* (Bi-Lateral Negotiation) is another system for negotiation training with virtual characters (Hill Jr., Belanich, Lane & Core, 2006). Like the TLTSs discussed above, it allows military personnel to negotiate with people from different cultures to train the former's interpersonal skills. One addition of *BiLAT* to these systems is the inclusion of a virtual tutor which provides hints and comments on the player's behaviour (Kim et al., 2009). Furthermore, it includes a module which provides a visualization of players' actions, and also gives feedback on how their actions influenced virtual characters, for example, by lowering trust, or improving rapport.

The above negotiation training systems focus on a domain that resembles that of law enforcement. Additionally, they tackle the training of interpersonal skills through different scenarios in which players interact with virtual characters. The TLTSs incorporate both verbal and non-verbal social behaviour by characters and enable the player to respond verbally. The characters update their negotiation strategies dynamically based on their trust in the players, which is influenced by their own responses throughout the narrative. In a retrospective on this work, Swartout (2010) notes that the characters, as well as the players, only have a limited amount of freedom. This is the case because the topics about which they can discuss are limited to the local context of the systems in such a way that players progress through dialogue trees.

5.1.2 *deLearyous*

DELEARYOUS Focusing on 'bad news' conversations between employees and managers, the serious game *deLearyous* models the interpersonal relations between the two interaction parties (Vaassen & Wauters, 2012). In *deLearyous*, virtual characters base their behaviour on their attitude towards the user. As a result, users (the managers) are required to learn how to behave to not let the virtual employees erupt in tears or anger.

INTERPERSONAL CIRCUMFLEX To model the conversations, *deLearyous* makes use of the theory of the *interpersonal circumplex* by Leary (1957). This model can be used to classify the attitude, or *stance*, of someone towards her interaction partner. The basic version of the interpersonal circumplex uses either 4 or 8 such different stances. In Section 5.2.1, I describe this model in more detail, as it is the same model we used in our *LOITER* games. Interaction in *deLearyous* transpires as follows. Users view a 3D virtual employee and can enter written text to speak to this person. This text is automatically interpreted in terms of stance and topic using machine learning. Automatic classification yielded an accuracy of 55% (with a random baseline of 24%) when using the interpersonal circumplex' 4 different stances. This

Figure 17.: An example visualization of the InterACT environment, showing the virtual human with four actions the player can take (adapted from Bosse and Provoost (2015)).

low accuracy is on par with human annotators, who rarely performed better at recognizing people's stances. Based on this finding, Vaassen et al. recommend that a tool such as deLearyous should always be used under supervision of a tutor to ensure that trainees learn the correct things.

5.1.3 STRESS

In the STRESS project, Bosse and Provoost (2015) investigate the development of INTERACT, an intelligent system which is able to analyse both decision-making process of humans in stressful situations and the causes of incorrect behaviour in such situations. Similar to the work I discuss in Section 5.2 and Chapters 6 and 7, the systems developed in STRESS deal with situations in which governmental personnel such as emergency workers or public transportation personnel have to handle social conflicts. Interaction with the virtual humans takes place in a virtual environment in which users can make choose between multiple options to take action, see Figure 17. An example scenario is that of interacting with a man who has not paid for his tram ticket and has to be convinced to exit the vehicle. The goal of the player is to determine which course of action results in a non-violent solution to the conflict. This depends on the personality of the virtual character, which can be detected from the type of aggression he applies: either reactive or progressive aggression. The available actions of the virtual character are influenced by his cognitive state. The user needs to learn to recognize the types of behaviours typify types of aggression and which actions work best to de-escalate the behaviour.

INTERACT

5.1.4 *FearNot!*

Already mentioned as one of the interactive storytelling systems in Section 2.1.3.1, *FearNot!* also deals with users being exposed to situations which involve conflict. In the *FearNot!* demonstrator, users do not directly influence characters' actions, but give advice on how they should act. By experiencing (second-hand) how this advice works out for such a character, users can reflect on whether their insights worked for that character. One of the capabilities of this system is its use of out-of-character reasoning (OOC), which lets characters choose to carry out an action for the highest dramatic effect. This way, when users give advice to a virtual character, that character and others interacting with her will exaggerate their possible responses, as they choose the actions that impact each other's feelings the most. Additionally, *FearNot!* also employs a *Story Facilitator* which chooses episodes for the story that match users' learning objectives (Aylett et al., 2008). Depending on the outcome of an episode, this module determines which conditions are met for successive episodes, for example, a character being in a certain emotional state or a change in relation between two characters. This way, *FearNot!* can adapt its story to learning objectives, yet its focus has thus far remained on selecting the most dramatically impacting episodes and events to happen.

5.1.5 *Nothing for Dinner*

NOTHING FOR
DINNER

5

IDTENSION

Nothing for Dinner is an interactive pedagogical drama in which players can experience what it is to live in a family of which one member has a traumatic brain injury, see Figure 18. The disruptive behaviour of traumatized individuals can affect family life, which constitutes the challenge for players of this simulation. Stepping away from the traditional distinction between strong autonomy and strong story approaches, *Nothing for Dinner* takes on an approach not unlike that of *Façade* (Habonneau, Richle, Szilas & Dumas, 2012).² It is based on *IDtension*, a narrative engine for interactive drama which models the narrative logic of a storyworld to generate appropriate actions (Szilas, 2007). Instead of only modelling the character actions, *IDtension* considers the narrative functions of those actions to create a proper narrative structure. This allows *Nothing for Dinner* to offer players various approaches to handle a situation, with characters responding in different ways.

5.2 LOITER

To design a serious game with which to improve social skills of police officers, I took three steps. Firstly, I looked at which theories can be used to describe social interaction and how we can use these theories to model such behaviour for virtual characters (Section 5.2.1). Secondly, I constructed a list of learning objectives that police students need to attain when learning about social interaction (Section 5.2.2). My third step was creating a prototype of a serious game called

² *Nothing for Dinner* is available at <http://nothingfordinner.org>.

Figure 18.: Player interaction with one of the characters of Nothing for Dinner (adapted from Habonneau, Richle, Szilas and Dumas (2012)).

LOITER to use as a test bed for techniques to improve social skills training (Section 5.2.3).

5.2.1 Modelling interpersonal behaviour

I collaborated with other researchers to determine which kinds of behaviour play a part in social interaction (Bruijnes, Linssen et al., 2015). We linked these behaviours to theories from social psychology that model them, based on annotation of police interviews. My research focuses on social skills in street interventions, which take place in the same domain, yet in a different context. Nonetheless, the role division and nature of conversations during both police interviews and street interventions are similar. Therefore, I base my further research for the behaviour of virtual characters in my research on this analysis of police interviews. I refer to Bruijnes (2016, pp. 87–111) for the complete analysis.

A police interview often is a situation involving conflict. Suspects might not cooperate with police officers and the police interview in general, but behave in a confronting manner. Suspects may be withdrawn, defiant or even aggressive towards officers. A police officer has the difficult task to convince the suspect to cooperate and tell the truth in an interview. We analysed the *Dutch Police Interview Training Corpus* (DPIT Corpus). This corpus contains interviews from 2012 and 2013 conducted by novice to moderately experienced trainees of the Dutch National Police. In 32 interviews with a total length of 13 hours, police trainees interviewed

DPIT CORPUS

professional actors based on 6 scenarios with different characters. The cast of characters included a young, naive girl, a man with an aggression problem and a criminal record, and a working class mother.

Figure 19 shows the steps we took in our analysis. In step 2, six observers independently selected fragments from the DPIT Corpus that they thought were ‘interesting’ in some way. For example, these were fragments in which a change in mood or atmosphere took place or fragments in which behaviour could be observed that indicated how the police officers or suspects felt about the interaction. Next, a subgroup of the observers noted as many different *terms* as possible to describe what was going on in the fragments (step 3). These could for example be adjectives describing the mood or nouns indicating behavioural traits. Based on these terms, we created questions with variations of the format ‘*To what extent is [term] the case?*’ (4). The original six observers then rated a random subset of the same fragments from the corpus on a five-point Likert scale for every question, once for the police interviewer and once for the suspect. For example, a question from the point of view of the suspect was ‘*To what extent is the suspect dominant?*’ (5). We performed a factor analysis to find a clustering of correlated questions (6). Next, a subgroup of the original six observers reached consensus on the interpretation of the factors (7) and we selected well known psychological and sociological theories that addressed these interpretations (8). We matched factors with the concepts underlying these theories (9 and 10). This also revealed the relations between the different concepts that were selected (11), as for some factors, concepts from different theories were applicable. This allowed us to integrate different theories into one ‘meta-theory’ that provides the terms and concepts to describe the interactions in a police interview. We checked whether this ‘meta-theory’ can describe what is happening in a police interview by analysing a second random subset of fragments from the corpus (12). In further studies, Bruijnes, op den Akker, Spitters, Sanders and Fu (2015) analysed such interviews in more detail with the goal of implementing a virtual suspect for simulation of a police interview (Bruijnes, op den Akker, Hartholt & Heylen, 2015) (13).³ Here, I briefly discuss the theories underlying the virtual suspect’s cognitive model.

In our analysis, we found that we could use a combination of three main theories to describe the interactions in the corpus on an affective level. These three theories are the following.

LEARY’S ROSE **INTERPERSONAL STANCE** Also known as *Leary’s Rose* and the *interpersonal circumplex*, this theory represents interpersonal behaviour in categories of interpersonal stance on the dimensions of affect (x-axis) and dominance (y-axis) (Leary, 1957; Gurtman, 2009), see Figure 20a. This circumplex can be divided in areas that represent are interpersonal stances, based on different degrees of affect and dominance. Furthermore, Leary (1957) suggests that human stances are affected by the interaction with the conversational partner. Through such *interpersonal reflexes*, conversational partners influence each other with their stances. Acts on the dominance dimension invite complementary behaviour while acts on the affect

3 For an overview of work on virtual suspects by Bruijnes, see Bruijnes (2016).

Figure 19.: Diagram showing the steps taken towards a computational model of police interviews (adapted from Bruijnes, Linssen et al. (2015)).

dimension invite symmetric behaviour (Figure 20b). For example, if someone displays dependent behaviour towards another person (submissive and positive), that other person will feel a tendency to adopt a leading stance (dominant and positive).

FACE AND POLITENESS The notion of a person’s public self-image, *face* (Goffman, 1959), influenced P. Brown and Levinson (1987) to construct their theory on *politeness strategies*. They distinguish between negative and positive face, which denote a person’s need for freedom (*autonomy*) and a person’s need to be approved of and approving of others (*approval*), respectively. Their approach to politeness hinges on the concept of face-threatening acts (FTAs) which are inherent with actions taken by a speaker, as these actions potentially impose on a hearer’s face by threatening their needs. For example, a person may request another to pass the salt, imposing on the other’s autonomy with her request. Brown and Levinson view politeness strategies as ways to redress these FTAs in order to minimise their imposition. These strategies range from being straight to the point, not wanting to hurt the other’s desires for autonomy and approval, to being indirect about one’s own needs. For example, when requesting salt, one might try to limit the imposition on the other’s autonomy by asking politely: ‘If you could, could you pass me the salt, please?’ Alternatively, a more direct approach is ‘Pass me the salt, please.’

FACE

FACE-
THREATENING
ACTS (FTA)

RAPPORT The feeling of *rapport* can be described as being ‘in sync’ with another person: communication takes place fluently and both interaction partners are roughly on the same level. Tickle-Degnen and Rosenthal (1990) conceptualised rapport in order to identify non-verbal correlates. Their description of the nature of rapport focuses on the interaction process as a whole and relies on three components of rapport: *mutual attention*, *positivity*

RAPPORT

Figure 20.: Leary's theory of interpersonal reflexes, also known as *Leary's Rose*.

Figure 21.: Importance of the three components of rapport over time (adapted from Tickle-Degnen and Rosenthal (1990)).

and *coordination*. Figure 21 shows a schematic view of relative importance of mutual attention and the other two factors of rapport over time. Mutual attention is beneficial to maintain a focused and cohesive interaction. Tickle-Degnen and Rosenthal (1990) mention that being positive towards each other is important during the build-up of rapport, yet becomes less important as time passes during interaction. Coordination is a harmonious relationship between partners. This is the key term related to the feeling of being 'in sync' and is the factor that becomes more important over time.

In addition to these three main theories to describe the interpersonal behaviour, we also constructed the meta-theories on exchanging information and framing the interaction, as well as on selecting strategies for goal-oriented behaviour (Bruijnes, Linssen et al., 2015). Interaction partners can deal with information in different ways, such as supplying it, requesting it, and expressing their psychological state

(Searle, 1969). Parts of a conversation can be framed so that it is clear to hearers and speakers what kind of behaviour is expected from them, for example when conducting small talk, telling stories or teasing people (Tannen, 1993). During communication, individuals make use of strategies to achieve their needs. Thomas (1992) argues that participants can take several approaches to resolve conflicts through communication, namely by accommodating to the other party, avoiding the conflict, competing with the other party, collaborating with the other party, and by compromising to find a solution to the conflict.

What we found in our analysis of the DPIT Corpus is that, to a large extent, interactions in police interviews are determined by the interpersonal stances of the parties. To get suspects to cooperate, it is often crucial to get them in a state of mind in which they are open to discussing and working out a solution together with officers. By being strategic about their approach, the officers can adopt a stance that invites suspects to cooperate. Then, officers can build rapport with suspects over time when they take people's face values into account. In such a way, police officers can ultimately gain the suspects' trust when they remain friendly. By either taking the friendly leading initiative in their interaction with the suspects, or by assuming a less dominant stance and asking them whether they would like to tell something, officers can resolve possible conflicts in interviews peacefully.

5.2.2 Learning objectives for social awareness

In the training curriculum of Dutch police trainees, there is no strict procedure for street intervention that trainees should master. Unlike in interviews with suspects or witnesses, there are no clear, context-independent rules for interactions with civilians. Instead, there is only a global guide line for police officers, namely to try to resolve situations by de-escalating possible conflicts and to only resort to more forceful means if the situation calls for it. Based on the analysis of police interviews we carried out (see Section 5.2.1) and further collaboration with the Dutch Police Academy, I created a set of learning objectives for social awareness of police officers to inform the design of LOITER. Before listing these objectives, I discuss the taxonomy I used to create and classify them.

Bloom's *Taxonomy of Learning Objectives* (TLO) is a classification of reasoning skills for use in class room settings (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956). It is meant to determine which kinds of cognitive skills students would require to attain given learning objectives. By classifying desired objectives in this taxonomy, teachers, instructors, and also game designers, can get a sense of which kinds of skills need to be trained in curricula they provide. Below, I describe the six different cognitive skill levels in order from low to high. The complete taxonomy consists of several subcategories for each of the six levels, except for *application*, yet I refrain from listing them here for simplicity's sake.

KNOWLEDGE Learning objectives which call for memorising and remembering facts, such as specific facts, conventions, and theories and structures.

COMPREHENSION Learning objectives which require students to interpret facts, and describe and discuss them.

APPLICATION Learning objectives which require students to apply their knowledge, for example, by solving a problem based on a memorised theory.

ANALYSIS Learning objectives which make students adopt a wider perspective by seeing patterns in the matter they are investigating, for example, by comparing a collection of theories.

SYNTHESIS Learning objectives which call for the creation of new theories or products, for instance, by predicting trends or by creating a game.

EVALUATION Learning objectives which require students to investigate information and criticise it according to some criteria, for example, by judging whether there is a bias in a research paper.

This ordering of levels of learning objectives implies a cumulative hierarchy: higher-order levels require understanding of lower levels. These six levels describe learning objectives with a *verb + noun* structure, for example, students need to *apply* their gained *knowledge*. In his overview of the revised version of Bloom's TLO (Anderson, Krathwohl & Bloom, 2001), Krathwohl (2002) explains how Anderson et al. incorporated a second dimension to distinguish between the types of knowledge the nouns refer to. This dimension distinguishes between the following four classes.

FACTUAL KNOWLEDGE Basic elements of a discipline, such as terminology.

CONCEPTUAL KNOWLEDGE Relations between the basic elements, for example, classifications and structures.

PROCEDURAL KNOWLEDGE Methods of carrying out something, such as specific algorithms or criteria to determine which procedures to use.

META-COGNITIVE KNOWLEDGE Students' awareness of their own knowledge, for example, being aware of which strategies they use.

I used this revised version of the TLO to construct a set of learning objectives that may be addressed in our serious games for social interaction (Linssen, 2014). In the previous section, I discussed several theories that give insight into interpersonal interaction. Rather than constructing learning objectives for all these theories, I decided to focus on Leary's theory of interpersonal stance (Leary, 1957). My reasons for this choice are two-fold. For one, I refrained from using all these theories to prevent the complete collection of learning objectives from being too complex for the students. Especially all the interrelationships between the theories might further convolute the students' understanding. Secondly, the instructors at the Police Academy base most of their curriculum on interpersonal skills on Leary's work on interpersonal stance. Therefore, it makes sense to support their current curriculum by matching our serious games to the training they provide. The theory about interpersonal stances captures the essentials of interpersonal interactions necessary to improve trainees' awareness, as it can be used to explain the important factors of different attitudes and willingness to cooperate.

Table 10.: The learning objectives for the theory on interpersonal stance for police trainees, classified in the revised version of Bloom’s Taxonomy of Learning Objectives by Krathwohl (2002).

KNOWLEDGE DIMENSION	COGNITIVE PROCESS DIMENSION					
	<i>Remember</i>	<i>Understand</i>	<i>Apply</i>	<i>Analyse</i>	<i>Evaluate</i>	<i>Create</i>
	<i>Factual</i>					
	<i>Conceptual</i>	LO1, LO5	LO2, LO3, LO6	LO4		
	<i>Procedural</i>	LO7		LO8 LO9		
	<i>Meta-cognitive</i>				LO10	

Generally speaking, there are three tiers of learning objectives trainees need to master. Firstly, trainees need to be able to recognise and classify behaviour in terms of stances and be able to select behaviour that expresses a specific stance. Secondly, they need to be able to describe how their own stance triggers the other person to adopt a certain stance. Trainees also need to be able to use this knowledge to their advantage by assuming a stance that triggers another person to behave according to a stance the trainee desires. Thirdly, trainees need to be able to break down and criticise an interaction. Below, I provide the complete list of learning objectives.

- LO1 Describe stances in terms of dominance and affect
- LO2 Describe stances in terms of the corresponding behaviour
- LO3 Recognise and classify behaviour in terms of stances
- LO4 Perform behaviour that conveys a stance
- LO5 Describe relations between stances
- LO6 Determine why an action evokes a particular response
- LO7 Know which stance to assume to invite a person to behave in a desired way
- LO8 Express a stance to invite a person to behave in a desired way
- LO9 Break down an interaction by determining points in time at which the stance of their interaction partner changed
- LO10 Criticise the approach they took in an interaction by explaining their choices, reflecting on it and indicating what they could have done better

Table 10 shows the classification of these learning objectives in the revised TLO. To improve the clarity of the two-dimensional TLO, Anderson et al. (2001) reworded Bloom’s original classes in the cognitive dimension to verbs, which I adopted in Table 10 as well.

I return to the above learning objectives in Chapters 6 and 7 to discuss how I investigated techniques to incorporate these objectives in our serious game. In the following section, I explain the design process of this game, LOITER.

5.2.3 Game design of LOITER

LOITER In collaboration with *T-Xchange*,⁴ I developed LOITER, a prototype of a serious game for social skills training. In this section, I discuss the design considerations for LOITER and describe a typical interaction with the game (Linssen, de Groot, Theune & Heylen, 2015). Based on this prototype, I implemented several techniques to support the learning objectives from the previous section. I discuss these implementations in Chapters 6 and 7.

5.2.3.1 Setting

DUTCH POLICE ACADEMY In the IUALL project I was involved in, I collaborated with the Dutch Police Academy (DPA) to inform the creation of our serious game. Their knowledge was particularly useful for the authoring of scenarios for LOITER. Based on examples from instructional material used by the DPA (Kaldenbach, 2006, 2008), I wrote several short scripts about possible street interventions. All these scripts involve a group of loitering juveniles who are causing nuisances throughout a fictional village. I chose these youngsters as the prime characters for LOITER because this would allow for a multitude of possible interactions. For example, the juveniles might be playing loud music in a calm neighbourhood, causing neighbours to complain, or they might be loitering in a business park after closing hours. By consulting with several police instructors, I created several settings for these stories. Equally important were the different approaches players would be able to take: these should mirror the actions actual police officers would take in real life. Before discussing the actual design and implementation of LOITER, I explain the considerations that played a role in creating this game.

5.2.3.2 Game style

While playing LOITER, players see the group of juveniles as shown in Figure 22. The background corresponds to a certain setting, while the main character with which players interact stands in front of the other two characters, who serve as supporting cast. A limitation on the interaction is that players only interact with this main character and not with the other characters in the group. This absolves the game of keeping track of the interrelations between the characters. These relations may play a role to some extent in real-life situations, but are not included in the current version of LOITER. Discussions with police instructors showed that interactions during street interventions often take place between a police officer and one individual from the group, who takes on the role of representing the group. In all the scenarios, this was 'Barrie', a 15-year old teenager.⁵

⁴ See <http://www.txchange.nl>.

⁵ The characters of LOITER lend their names from the Dutch TV series *New Kids*.

Figure 22.: The three juveniles of LOITER, loitering in a shopping street.

One of my assumptions in creating LOITER was that a completely realistic simulation of a social interaction does not necessarily lead to effective learning (Linssen, de Groot, Theune & Bruijnes, 2014). Alexander, Brunyé, Sidman and Weil (2005, p. 362) assert that ‘if the level of fidelity captures the critical elements/properties of the skills/tasks you wish to teach, that level of fidelity is sufficient even if it noticeably deviates from the real world.’ For LOITER, the learning objectives described above emphasize gaining insight into the workings of Leary’s interpersonal circumplex. Therefore, I decided to provide players with a game that focuses on the higher-level thought processes of virtual characters that govern their actions, instead of placing emphasis on nuances in verbal and non-verbal behaviour. While these undoubtedly play a role in social interaction and students should train their recognition abilities of such nuances, LOITER’s learning objectives are on the more conceptual level of the theory on interpersonal stances, as shown in the previous section.

5.2.3.3 Behaviour of the virtual juveniles

The actions of characters in LOITER are based on their stance towards the player, which may change over time due to the player’s actions. Like the VST, LOITER uses a turn-based mode for interaction. Players interact with the main character, Barrie, in a multiple-choice fashion. Each of the options for behaviour from which players can choose relates to a particular stance from the interpersonal circumplex. To prevent an overload of choices for players, LOITER uses four stances, built from combinations of high and low dominance and affect. These are, in clockwise order starting from the top-left: *attacking*, *leading*, *accommodating*, and *aloof*. Figure 23 shows examples of these four stances in the way they are presented to players in LOITER. To give players sufficient insight into how their action is uttered, the verbal utterances are paired with non-verbal postures of the officer, which resolves possible ambiguity of the verbal utterances. These postures are based on previous

Figure 23.: Four different actions available to players in LOITER, corresponding to four different stances. Starting from top-left in a clockwise order: attacking, accommodating, aloof, leading.

research on the relations between non-verbal behaviour and interpersonal stance (Ravenet, Ochs & Pelachaud, 2013). While playing the game, available actions are shown to players in random order, so that players need to read the utterances and interpret them before making a choice and not simply select an action based on its position on the screen.

Barrie, the main juvenile, responds both verbally and non-verbally, based on a player's action. Similar to the four actions available to players in their turn, Barrie has four responses corresponding to the four stances in the interpersonal circumplex. To prevent a combinatorial explosion of possible action-response pairs, often experienced while designing branching narratives, I authored the scenarios in such a way that any of the four responses of the juvenile can follow every of the four actions of a player. Barrie chooses to carry out the action that corresponds to his current stance. This stance is influenced by a player's action, following the theory of Leary's interpersonal reflexes (Leary, 1957). For example, this means that when a player chooses an *accommodating* action, the juvenile will be invited to perform a *leading* action and vice versa. To prevent the juvenile from very frequently and suddenly changing stance when players decide to take different actions, and coming across as inconsistent, his stance does not necessarily change with every action a player takes.⁶ The juvenile's stance is modelled in a two-dimensional grid, see Figure 24. Players' actions correspond to certain positions in the two-dimensional grid as indicated in Figure 24. Barrie's stance can also be represented in that grid; initially, he will have a neutral stance. As in the example above, when a player chooses to act *accommodating*, her action will invite an opposing stance, namely a *leading* one. Barrie's stance will change based on this action by moving towards the invited stance. Both dimensions in this model range from -2.5 to 2.5 . The change of Barrie's stance is done with a fixed step-size of 1 (in one or both dimensions), so that it may take players multiple turns to get the juvenile in the desired quarter. Barrie will respond according to a certain stance when his position in the grid falls within the boundaries of that stance. For example, for the *leading* stance, this means that his affect has to be larger than 0 and his dominance also larger than 0. This way, Barrie will only switch stance when he crosses such a boundary. For example, when a player has taken an *accommodating*

⁶ However, this does not put limitations on players' behaviour: if they choose to switch stance often, they might come across as inconsistent to the other party as well.

Figure 24.: The two-dimensional grid in which a character's stance is modelled in LOITER. The player's actions are indicated by $P_{\text{attacking}}$, P_{leading} , $P_{\text{accommodating}}$, and P_{aloof} . Barrie's initial stance is labelled as B_1 and his stance resulting from the player's $P_{\text{accommodating}}$ action is labelled B_2 .

action ($P_{\text{accommodating}}$), she invites Barrie to take a leading stance. Initially, Barrie's stance is neutral (0,0) and, after the player's accommodating action, his stance shifts with the fixed step size to (1,1), as shown in Figure 24. When the player has taken multiple accommodating actions, Barrie's position in the grid will eventually reach the limit of (2.5,2.5). As a result, it will take that player more turns to get the juvenile in a different stance, when she performs other actions. This shows that there are long-term effects of consistently acting in a particular way, as opposed to immediately getting a response to one's action according to the theory on interpersonal reflexes.

5.2.3.4 Scenario

The scenario for the first iteration of LOITER involved the group of loitering juveniles playing loud music in an otherwise peaceful neighbourhood. This scenario includes a description of the situation, an introduction for players and all actions and responses of both the player and the juvenile. A police instructor from the DPA provided minor comments and suggestions on the actions in the scenario. A typical playthrough session with LOITER goes as follows. When starting the game, the interface shows an introductory text, explaining the setting and the in-game goal, see Figure 25. This in-game goal is for the player to resolve this situation in a peaceful manner. Furthermore, the text shows players how to interact with the loitering juveniles and on what information she should base her choices. Specifically, the text explains the basics of Leary's Rose as well as the fact

Figure 25.: The introduction screen of LOITER, showing text explaining the game's setting and mechanics.

that the available actions and responses relate to that theory. However, it does not explain how these actions and responses correspond to Leary's Rose, as this is for the player to figure out while playing the game.

The player continues to the interaction with the juveniles, which looks like Figure 26. Here, the juvenile's non-verbal posture is shown together with the player's available choices, in random order.

When the player chooses to carry out an action, her choice is visualized in the interface as shown in Figure 27, in this case, an *aloof* approach. The game waits for the player to indicate that she has seen the action she chose before showing the juvenile's response: an *attacking* reaction (Figure 28). Barrie changes his posture according to the representation of his stance in the two-dimensional grid. This always happens in the first response, as Barrie's stance coordinates shift to one of the four possible stances. In subsequent action-response pairs, Barrie's stance may change or not, depending on the actions the player chooses, as described above.

In total, the player takes eight actions throughout playing this scenario, witnessing as many responses from the juvenile. The scenario is divided into the following four parts, roughly following a story arc like that of Freytag's Pyramid (see Section 2.1.3.2). Each part contains two action-response pairs.

THE APPROACH The player indicates how they approach the juveniles: dominant and direct, or in a more easy-going way. Barrie asks them why they are here.

Figure 26.: The interface of LOITER when players can make a choice in how to act.

5

Figure 27.: The interface of LOITER when players have chosen an action.

Figure 28.: The interface of LOITER when the juvenile responds.

5

THE PROBLEM The player explains the reason they are here is the music the juveniles are playing, with a disgusted or understanding tone. The juvenile asserts that they like listening to their music this way.

THE SOLUTION The player indicates that there is a neighbourhood centre nearby where the juveniles can relax the way they want to. Barrie expresses his hesitation or good faith in this plan, depending on the players' behaviour.

THE ENDING The player leaves with a message ranging from a threat to a reassurance that things will be all right. The juveniles take their leave with a grudge or a convinced attitude.

On completion of the scenario, LOITER presents a GAME OVER screen to the players with feedback on which stance they most frequently assumed, as well as which stance Barrie most frequently enacted.

5.2.3.5 Relation to the VST

Originally, LOITER was meant to be a successor to the Virtual Storyteller (VST), inheriting the latter's architecture and emphasis on an emergent narrative (see Section 2.2). However, the domain of law enforcement and specifically the setting of street interventions rely heavily on dialogues. The VST's existing architecture did not support dialogues particularly well, which called for a revision of how such interactions are handled. Currently, the VST treats speech acts and other actions similarly: given the right preconditions, actions may be triggered, resulting in new

events and facts becoming known. For dialogues, and especially the modelling of the cognitive states of the characters, this would require keeping track of several parameters, such as known information and characters' attitudes towards other characters. Instead of addressing these challenges in the VST's framework, I decided to create a prototype which could be managed and authored in a manner that involved less effort, allowing me to make alterations to this prototype more rapidly. Moreover, this allowed me to focus on the design of techniques to improve learning through this serious game.

LOITER is a spiritual successor to the VST. It takes a character-centric approach to storytelling, however, it does not allow the characters much freedom. Nonetheless, characters in LOITER base their actions on their personal cognitive states. Players' freedom is not as large as in the VST either, as the game is authored in such a way that there is one main storyline that can be explored. In Chapters 6 and 7, I discuss the links to the VST in more detail.

We don't need other worlds. We need mirrors.

Snaut
Solaris
Lem (2011)

In this chapter, I discuss how the meta matters to players of serious games.¹ This chapter addresses Research Question 3 of this dissertation:

RQ3 *How can serious games stimulate meta-thinking to increase their effectiveness?*

To answer this research question, I investigate feedback techniques aimed at stimulating reflection in players. In the previous chapter, I discussed LOITER, a serious game aimed at improving social awareness of police trainees. Players play this game by conversing with a virtual character: a loitering juvenile who is causing a nuisance together with some friends. The player's goal in the game is to prevent this possible conflict from escalating by choosing actions that invite the juvenile to cooperate. The overarching learning goal of this game is that players need to become aware of the effects their behaviour has on that of the virtual juvenile (see Section 5.2.2 for the specific learning objectives). To instil this awareness, I implemented two *meta-techniques* in LOITER, based on techniques used in *live action role play (larp)*. These techniques allow for the communication of information which is normally not available to players in-character, such as the thoughts of other characters. As such, these meta-techniques are an out-of-character means of communication and allow roleplayers to go meta to enrich the experience of a larp. In this chapter I explain how virtual agents can present information about their thoughts by going meta, showing their thoughts in ways similar to those found in larps. This meta-information allows a player to get a better insight into the causal relations between her actions and the virtual juvenile's reactions.

In Section 6.1, I discuss the possibilities for support through feedback in serious games. In Section 6.2, I explain the basis of the design of the two meta-techniques I implemented in LOITER: thought bubbles and flashbacks. I explain the implementation of thought bubbles in Section 6.3 and discuss its evaluation with police trainees in Section 6.4. Similarly, I explain the implementation of flashbacks in Section 6.5 and its evaluation in Section 6.6. I close this chapter by reflecting on the research question, discussing how meta-techniques can be used to stimulate reflection in serious games in Section 6.7.

META-
TECHNIQUE
LIVE ACTION
ROLE PLAY
(LARP)

6

¹ This chapter is based on Linssen, Theune and de Groot (2013), Linssen (2014), and Linssen, Theune, de Groot and Heylen (2015).

6.1 THEORIES AND IMPLEMENTATIONS OF FEEDBACK

Feedback is present in all games. Players of games take actions which result in observable effects. For example, walking to a different location reveals that location, talking to someone makes them talk back, or clicking on a button performs an action in the game. The feedback I discuss in this chapter goes beyond this directly observable form of feedback. The crux of serious games is that people learn from their in-game behaviour, which can be supported by reflecting on it, as Kiili (2007) and Wouters and van Oostendorp (2013) argue. In this context, feedback is meant to elucidate *how* and *why* certain actions in a game have certain outcomes and, moreover, how this extends to real-life knowledge. In this section, I first present a literature study on how feedback can be presented in serious games, focusing on games for social interaction. Secondly, I discuss several implementations of feedback in such serious games.

6.1.1 Support through feedback

LAW OF
EFFECT

FEEDBACK IN-
TERVENTION
THEORY

FORMATIVE
FEEDBACK

The use of feedback as a way to improve a trainee’s performance has spawned a long line of research (Mory, 2004). This emerged from Thorndike’s *Law of Effect*, which asserted that feedback is a ‘connector’ between cause and effect (Kulhavy & Wager, 1993). Similarly, Kluger and DeNisi (1996) describes *Feedback Intervention Theory* which can integrate the previously existing theories concerning feedback and other fields, such as those on task-related learning and motivation. Feedback can be used to explain events and correct trainees. As the field evolved, it diverged into different explorations of the mechanics underlying proper use of feedback. Shute investigated *formative feedback*, defined as ‘information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning’ (Shute, 2008, p. 154). She stipulates three key elements of good formative feedback: usefulness to the learner, availability during a learning task, and motivation on the learner’s behalf. Dunwell, Freitas and Jarvis (2011) provide a guide to design feedback, using a four-dimensional approach comprising learner characteristics, learning requirements, learning context and available technology. These four dimensions assist in deciding the type of feedback, its content, format and frequency. For example, feedback may be evaluatory, informing players how well they did using emojis at the end of the game. Feedback may also be interpretative, for example, in the form of text messages explaining how well players performed and why they were assessed as such.

In commercial off-the-shelf games, many game mechanics deliver an explicit form of feedback. One example is a numerical score which keeps track of a player’s health, showing the player how many enemy hits she can still endure before her player character succumbs. Another example is the indication of the fact that the player’s choices have an effect on non-player characters in games. This is done in Telltale’s *The Walking Dead* video games (Telltale Games, 2012), which inform the player, through on-screen text, that a certain character may remember a decisive action of the player. In the context of serious games, many forms of feedback for educational purposes have been explored in relation to game mechanics (Arnab

et al., 2014; Kiili, 2007). One typical type of feedback is the *after-action review*, which provides feedback on a player's performance when she has completed a game scenario, see, for example, Lane, Core and Lent (2005). There also are tutoring coaches which accompany a player during a scenario, providing intermittent feedback while the game is played, see, for example, Winer et al. (2012). These approaches offer *metacognitive* feedback: they are intended to let players explicitly reflect on their in-game behaviour. This is meant to instil awareness in them about the effects of their behaviour on the virtual world and how they might have acted differently. Moreover, the insights based on this feedback should be carried over from the virtual world to the real world. As I also discussed in Section 1.2, this concept is exemplified in the Serious Game Lemniscate Model (SGLM), which describes how in-game behaviour can be reflected upon. This takes the form of two cycles, as shown in Figure 29. In the left cycle, players plan and take actions in the virtual world based on the input they receive from interacting with it. In the right cycle, players *go meta* and internalize their experience in the virtual world, reflecting on how it applies to them in the real world.

AFTER
ACTION
REVIEW (AAR)

Figure 29.: The Serious Games Lemniscate Model (adapted from Koops and Hoevenaar (2012)).

From the existing research on feedback, I draw two main lessons. Firstly, it is vital to the design of feedback that it aids in the learning process, making it more effective or more efficient. This should be done by making players aware of how their in-game behaviour relates to the real world. Secondly, feedback needs to be tailored to the gaming context. It is important that the feedback does not hamper the gaming experience and is delivered in a timely manner, depending on its information and the task at hand in the game. Before addressing how I designed several techniques to provide feedback in LOITER, I discuss existing implementations of feedback in serious games.

6.1.2 Implementations of feedback in serious games

TUTORING	<p>As I explained above, feedback can have various functions and take on many forms, in different combinations. In serious games for training social skills, one typical form of feedback is the after-action review (AAR), which is used by several TLTS which were discussed in Section 5.1. In these systems, trainees receive overviews of their interactions with virtual agents in various manifestations, such as text summaries or recorded video material showing their in-game actions. These overviews can then be used for discussion with peers and tutors. Core, Traum, Lane and Swartout (2006) have augmented these AARs with additional information about the virtual agent's thought processes. This information can be shown next to a summary of the interaction, indicating how the relationship between the trainee and agent has been influenced. Alternatively, players can question the virtual agent about its feelings at a certain moment during the scenario, letting it explain how the player's actions influenced its cognitive state. A similar approach is found in <i>BiLAT</i> (Kim et al., 2009), in which a virtual agent tutors trainees by providing a walkthrough of the scenario and explaining how the player influenced the virtual agent's behaviour in the scenario. This is an example of <i>tutoring</i> which explains concepts related to the task at hand or offers more information about the causes and effects of events in a scenario (van Lehn, 2011). <i>BiLAT</i> was evaluated by having trainees complete situational judgement tests (SJTs), which consisted of short descriptions of scenarios and possible responses, and comparing the trainees' ratings of the effect of these responses to the ratings of subject matter experts. For trainees without prior negotiation experience, their experience with <i>BiLAT</i> resulted ratings closer to those of the experts than before (Kim et al., 2009).</p>
SITUATIONAL JUDGEMENT TEST (SJT)	<p><i>DynaLearn</i> offers support by tutoring agents which provide tips and hints to players while they are playing a scenario, as opposed to providing this information afterwards (Winer et al., 2012). Moreover, the tutors employed by this system have different personalities that match the type of feedback they offer. For example, the 'Mechanic' assists trainees in analysing their model and the 'Quizmaster' can question trainees using short quizzes to assess their current knowledge. Conversely, <i>FearNot!</i> (see Section 2.1.3.1) switches these roles: students take on the role of a tutor to give feedback to a virtual agent on how to cope with a bully (Aylett et al., 2005). On the whole, <i>FearNot!</i> improved children's cognitive empathy, yet it is unclear to which mechanics of the system this can be attributed (Vannini et al., 2011). Broekens et al. (2012) developed a virtual reality environment for training negotiating terms of employment with virtual agents. Their approach makes use of thought bubbles of the virtual agents to show preferred types of explanation to trainees without disrupting the flow of the interaction with the agents. However, they do not report the effects of this feedback, if any, on the learning outcomes of users of this environment. Allain and Szilas (2012) discuss the use of <i>metalepsis</i> by virtual agents in the context of serious games: they address the player in an out-of-character way, expressing their opinion about the player's behaviour. This is an alternative to feedback presented in classical user interface elements, intended to be a less intrusive way of delivering such information as compared to classical interface elements.</p>
METALEPSIS	

While serious games do include feedback mechanics for deeper reflection on the user's experience, these implementations of feedback are not investigated in isolation. Therefore, it is unclear to what extent feedback supported the learning experience in these serious games. In this chapter, I address this lack of research on feedback mechanisms by comparing different versions of our serious game, LOITER, which incorporate or lack feedback.

6.2 META-TECHNIQUES IN LOITER

To design feedback mechanisms for LOITER, I turned towards the paradigm of *going meta* once more. From the theories on feedback, I conclude that explicit reflection is a way of going meta: trainees are required to reason about their experience to learn from it. Another insight came when I delved into the origins of the Virtual Storyteller, which can be found in improv theatre, which in turn is very similar to *larps*. Below, I describe how techniques used to improve experiences of *larpers* (people engaged in a *larp*) relate to the design of the techniques I implemented in LOITER to improve reflection.

6.2.1 Live action role play

Larping is a form of role playing in which participants physically enact their character roles. Especially in the Nordic countries of Europe, this form of role playing has a large following and the Nordic *larp* tradition has taken on a multitude of forms (Belarbi et al., 2010). These varieties of *larp* range from simulations in which every player tries to strictly stay in character to *larps* that are controlled by a director who can instruct characters to behave in certain ways during the play. In the latter, the distinction between the story world and the meta-story world is exploited to let the director influence how the play turns out. This is an example of a *meta-technique* *larpers* communicate information between each other that would not have been available to them in character. Meta-techniques can thus be used in *larp* as a communication channel to direct players, but also to enrich the story and characters' backgrounds, for example through flashbacks or inner monologues. A list of several meta-techniques taken from *larps* follows below.

BLACK BOX An area that is shielded from a *larp*'s common role-playing space, and which can be used by players to use various meta-techniques in a non-intrusive manner, that is, without disturbing others' play. The black box is mostly used to enact scenes that do not fit into the *larp*'s time/space continuum, for example, flashbacks.

FLASHBACK AND FLASH-FORWARD Flashbacks may be used by players to enact scenes that happened in the past (before the *larp*'s time frame) or may happen in the future.

INNER MONOLOGUE Players speak out their characters' inner thoughts aloud for other players to hear. This also gives insight into characters' motivations.

ACT BREAK It may be agreed upon by larpers to break the larp into several acts. In between these acts, act breaks can be used to let players discuss the events that occurred during the previous act and how things may or should turn out during the next. This the larp equivalent of an after-action review as discussed in the previous section.

META HOUR Similar to act breaks, the meta hour is a meta-technique during which the activity in the larp is paused. Then, scenes that took place during the larp can be re-enacted in isolation so that all players can witness these events.

RUN-TIME GAME MASTERING A larp may include players that do not act out character roles but instead direct or 'master' the larp. This can happen in a variety of ways, for example by 'shadowing' players and acting as their inner voice, speaking aloud the players' thoughts so that they must act accordingly. Masters may also manipulate the environment or even the player directly, for instance by shifting around objects in the world or by steering a player toward a relevant position in the world. Such a game master takes on the role of a Director Agent as in, for example, the Virtual Storyteller, or the Story Facilitator in FearNot!

6.2.2 *Feedback through meta-techniques*

Meta-techniques are used in larps to deepen the background stories of characters or show possible consequences of larpers' choices. We can incorporate these meta-techniques in games to achieve social believability in similar ways. This may have several interrelated effects. An increase in social believability may influence players' feelings of both presence and immersion. That is, because the characters in the game are more believable, it may be easier for players to feel as if they are present in that virtual environment and, subsequently, it may improve their immersion (Jennett, Cox & Cairns, 2008). When the purpose of the game is to simulate social situations (as in LOITER), increased social believability of the characters is likely to increase the players' degree of learning from their social interactions in the game. When players feel like they are present in the virtual environment and are immersed in the game, the learning effect may be even greater, as the social situations will seem more real to players.

Rather than addressing the challenging task of letting characters display social signals in-character (Vinciarelli et al., 2012), meta-techniques may be used to convey their mental states out-of-character. This can be seen as an extension of 'explainable artificial intelligence' (van Lent, Rey, Fisher & Mancuso, 2004) as characters will be able to explain to players how actions and events affected their own behaviour. For example, the use of flashbacks can give insight in the emotional states and social attitudes of a character. The inner monologue meta-technique can also be implemented as a game mechanic that explains what is going on in the character's mind without having to actually show it in character. *Carmen's Bright IDEAS* uses a variation on this theme to provide players with choices in its interactive experience (Marsella, Johnson & LaBore, 2000). In this multimedia title,

players adopt the role of the mother of Carmen, a teenage girl. Thought bubbles are shown to allow players to choose the mental state of Carmen, influencing how the story progresses. Another use of meta-techniques can be found in virtual agents that act as a couple consisting of a devil and an angel could give advice to players on certain situations that involve social dilemmas. This pair of agents could then function similarly to tutoring agents in that they give suggestions to the player, spelling out what players' options could be. For example, in the case of (criminal) youths, the couple could mention two alternative choices: giving in to peer pressure or obeying the demands of a police officer. Act breaks and meta hours can be used in games to actively discuss what is going on in inside the characters' minds, as a form of AAR. Such meta-techniques allow players to reflect on the emotional and social motivations for both their own and the virtual characters' actions.

As discussed above, meta-techniques are usually used to communicate information on a 'player' level; in other words, it happens out-of-character. For example, an act break might be incorporated during the game as a questionnaire about players' feelings about other characters. Because this questionnaire is not part of the story world, players are forcefully cast out of their character role. This is likely to reduce the players' feelings of presence in the story world and their immersion in the game. Both Jennett et al. and Lukka point out that such switching between IC and ooc roles may lead to confusion and may disrupt immersion, so this should be taken into account when designing meta-techniques as feedback mechanisms (Jennett et al., 2008; Lukka, 2011). Yet there are different ways to implement meta-techniques, with some being less intrusive than others. It seems that for many intrusive techniques, which take players out of their character roles and thus also out of the story world, non-intrusive counterparts can be found that allow the player to stay immersed within the story world as much as possible. For example, a less intrusive employment of the act break meta-technique would be to include scenes within the story in which players and NPCs discuss their relationships to other characters. These scenes could for instance take the form of in-character sessions with a therapist. Another non-intrusive meta-technique would be implementing the angel/devil couple as actual characters in the game who give advice to players. The inner monologue technique could simply be implemented as a character that is talking to itself, which just might happen to be overheard by other characters.

In short, a concern with intrusive implementations of meta-techniques is that they may lead to a decrease in presence and immersion, while non-intrusive implementations do not suffer from these effects. However, it is not *a priori* clear how the reduced immersion caused by intrusive meta-techniques will affect learning. As I re-iterated in the previous section, the SGLM emphasizes that, at key points during gameplay, players should be 'pulled' out of their character frame to actively reflect on what has happened (Koops & Hoevenaar, 2012). Indeed, this model advocates intrusive use of such learning moments to make players aware of how their attained knowledge relates to the world outside of the game. In other words, intruding on the player's feelings of presence and immersion may lead to better learning than when reflection is not that explicit. Yet this may not necessarily be the case for games aimed at improving social awareness, because

becoming aware of social behaviour may be done best by feeling as present in the virtual world as possible. So it seems that, in the context of social awareness games, the intrusive use of meta-techniques may have both positive and negative effects on learning. Thus, while the way in which meta-techniques should be implemented is clearly dependent on the context and the purpose of games, the question remains how these implementations should be carried out to allow for the most enjoyable or efficient (in terms of learning) experience. In the following sections, I describe how we addressed this in the design and evaluation of two meta-techniques in LOITER: thought bubbles and flashbacks.

6.3 THOUGHT BUBBLES

LOITER-TB

In the previous section, I defined the key functionality of meta-techniques in serious games as providing feedback to users to aid in their learning process. In this section, I discuss the first implementation of a meta-technique in LOITER, namely that of thought bubbles. We designed thought bubbles to fit into the virtual environment of LOITER, creating a new prototype we call LOITER-TB (*Thought Bubbles*). In LOITER-TB, we use thought bubbles to help players achieve the learning goals from Section 5.2.3 by providing short-term feedback. For LOITER-TB, we chose to include short moments for explicit reflection so that players can go meta without letting this technique be too intrusive to their gaming experience.

The meta-technique on which we based our use of thought bubbles is that of the inner monologue, as explained in the previous section. In larps, larpers can use this technique to narrate their thoughts so that other players may become aware of them. In LOITER-TB, we provide players with extra information in thought bubbles about the cognitive state of the virtual characters they interact with to help them attain the learning goals, in a way similar to that of Broekens et al. (2012).

LOITER-TB is based on LOITER as described in Section 5.2.3, featuring a short scenario that revolves around loitering juveniles playing loud music. The player's goal is to end this nuisance. He or she can do so by choosing from four different options on how to respond both verbally and non-verbally to the leader of the juveniles, based on the four stances from Leary's Rose. The player's choices influence how the juveniles will leave, based on the 'inviting' relations from Leary's model. For example, the juveniles may do so willingly or grudgingly, with the implication that the way in which they leave has an effect on future encounters. The postures of the virtual characters are based on previous research on non-verbal expression of interpersonal stances (Ravenet et al., 2013). Figure 30a shows the interface of the game with a player action, the juvenile's reaction and his thoughts about the player's stance.

For the content of the thought bubbles, we used Dutch adjectives that correspond to stances in Leary's Rose, following research by Rouckhout and Schacht (2000). Based on interviews with instructors from the Dutch Police Academy, we chose to include mnemonic pictures of animals that are archetypical for the different stances, as they are also used in this way in the current training curriculum of police officers, see Figure 30b. Players receive a short instruction on the meaning of the thought bubbles before playing through the scenario, see

Figure 30.: The initial design for feedback through thought bubbles in LOITER-TB. (a) A screenshot of LOITER-TB when the player has just made a choice. (b) The visualizations of the four stances in the thought bubbles.

Figure 31. On finishing the game, players receive a short debriefing (a condensed AAR) of what happened in the scenario in which they are shown the result of their actions, see Figure 32. The debriefing screen also indicates the stances both the player and the leader of the juveniles assumed most frequently. This allows for further reflection of players on their actions.

6.4 EXPERIMENT: THOUGHT BUBBLES IN LOITER-TB

To evaluate our implementation of thought bubbles, we conducted an experiment with police trainees. Because meta-techniques provide insight in a character's thoughts, we expect that the information presented in thought bubbles aids trainees in learning about social interaction, as compared to a version of the game without the thought bubbles. We address the following three hypotheses.

HYPOTHESIS 1 Thought bubbles help trainees to attain knowledge about social interaction that is similar to that of domain experts.

HYPOTHESIS 2 Thought bubbles help trainees to determine which kinds of behaviour their own behaviour invites.

HYPOTHESIS 3 LOITER-TB helps trainees to recognize stances of other people.

6.4.1 Method

We conducted an experiment with a repeated measures, between-subjects design with two conditions. Participants in condition TB played LOITER-TB with thought bubbles. Participants in the control condition (c) played the game without thought bubbles. The game featured eight choice points at which players chose between behaviours corresponding to the four stances (see Section 5.2.3).

We used two measures to determine the participants' knowledge about social interaction in street intervention: Pathfinder network analyses and *situational judge-*

Figure 31.: The explanation of thought bubbles in LOITER-TB.

Figure 32.: The condensed after-action review in LOITER-TB.

Figure 33.: The averaged Pathfinder network of the four domain experts. Edges in the graph indicate strongly related concepts.

ment tests (SJTs), similar to *BiLAT* (see Section 6.1.2). A Pathfinder network (*PFnet*) analysis can be used to calculate a person's mental model based on ratings of the relatedness of given concepts (Wouters, Oostendorp, Boonekamp & van der Spek, 2011). For this analysis, we first created a set of ten concepts important to social interaction in collaboration with police instructors. These are, in alphabetical order: attacking, aloof, accommodating, helpful, de-escalation, escalation, giving freedom, giving respect, working against each other, and working together. Four domain experts (one police instructor and three researchers with expertise on Leary's Rose) rated the relatedness between these concepts in 45 concept pairs on a 7-point Likert scale, which ran from 'completely unrelated' to 'extremely related'.² Figure 33 shows a *PFnet* of the averaged ratings of the domain experts. Links in this *PFnet* indicate strong relations between the linked concepts. By comparing the participants' *PFnets* to the averaged *PFnet* of the domain experts, we analysed to which extent the *PFnets* were similar. An additional indicator of similarity of mental models is the coherence of *PFnets*. This is an indirect computation of the similarity between concepts, indicating whether concepts that are rated as related have similar links to other concepts (see Wouters et al. (2011, p. 745)). We address Hypothesis 1 by comparing the *PFnets* of participants to those of domain experts.

PFNET

As the second measure, participants took a situational judgement test (SJT) consisting of two parts. In the first part, participants indicated which stance a certain (visually displayed) behaviour of the juvenile reflects. For example, the following utterance of a juvenile was shown, together with the corresponding non-verbal posture, meant to convey a submissive-friendly stance: 'OK OK, we'll do as you say.' In the second part, participants were shown an utterance of the police officer, based on one of the four stances, for example, 'We can easily solve this together!' to express a dominant-friendly stance. The participants indicated which of the four stances they would expect the juvenile to adopt in response to the police officer's utterance. We used these SJTs to assess participants' abilities to determine

² The number of concept pairs that should be rated is calculated as follows: $\frac{n(n-1)}{2}$, with n the number of concepts.

invited behaviour and to recognise stances of others, addressing Hypotheses 2 and 3. Appendix B.1 lists both SJTs.

The experiment followed a pre-test, activity, post-test structure. Firstly, participants rated the 45 concept pairs to build a PFnet. Secondly, they took the SJTs. Having completed these pre-tests, participants were instructed about the scenario and game mechanics of LOITER-TB through the in-game introductory text. They played the game twice, allowing them to become more familiar with the scenario and the gameplay, and to experiment with different behaviours. Having finished the game, participants took the post-tests: the same PFnet test and the SJTs as in the pre-test. Participants also rated their experience of the game on nine 7-point Likert items. For example, participants indicated how much they agreed with the statement 'The scenario played out the way I expected it to do.' Appendix B.1 lists the complete game experience questionnaire. The experiment concluded with a demographics questionnaire in which participants were asked to report their gender, age, and experience with computer games in general.

We conducted the experiment with a class of 21 third-year police students of the Dutch Police Academy, who were being trained to become senior police officers. These students already had practical experience by accompanying graduated police officers on the streets. The mean age of the participants was 33.0 ($SD = 6.9$); 71% was male and 29% female. On average, it took participants 30 minutes to complete the experiment, of which 6 were spent on playing the game. From the 21 participants, 12 were included in condition TB and 11 in condition c. Of the latter, two participants experienced technical issues, leaving only 9 participants of which the results are taken into account for analysis.

6.4.2 Results

We compared the participants' pre- and post-test PFnets to an average of the PFnets of the four domain experts. Using two ANCOVAs (dependent variables: post-test coherence and similarity scores, fixed factor: condition, covariates: pre-test coherence and similarity scores, respectively), we did not find significant differences between conditions for the coherence of the PFnets, nor for the similarity between participants' and domain experts' PFnets ($F(20) < 1, p > .05$). A paired-samples t-test showed that the coherence of participants' PFnets in both conditions increased significantly ($t(20) = 3.329, p = .003$; pre-game: $M = .370, SD = .363$; post-game: $M = .584, SD = .338$). Another paired-samples t-test did not show a significant difference ($t(20) < 1, p > .05$) between the pre- and post-game similarity between participants' and domain experts' PFnets across conditions.

An ANCOVA (dependent variable: post-game test score, fixed factor: condition, covariate: pre-game test score) did not show a significant difference ($F(20) < 1, p > .05$) between conditions for the participants' ability to indicate invited reactions of a juvenile (TB: $M = 2.420, SD = 1.084$; c: $M = 1.67, SD = .866$).

A paired-samples t-test showed no significant change ($t(20) < 1, p > .05$) in the participants' ability to recognise the juvenile's stances before ($M = 4.81, SD = .981$; correctly answered questions, out of 8) and after ($M = 5.05, SD = 1.284$)

the game was played. An ANCOVA (dependent variable: post-game test score, fixed factor: condition, covariate: pre-game test score) also failed to show a significant difference ($F(20) < 1, p > .05$) between conditions for the participants' ability to recognise a juvenile's stance.

6.4.3 Discussion

Based on the PFnet analyses, Hypothesis 1 cannot be proved. We conclude that LOITER-TB aids participants in constructing more coherent mental models, but not in constructing mental models that are more similar to those of domain experts. The addition of thought bubbles in LOITER-TB does not improve the construction of mental models. Because the participants were already experienced at street interventions, we believe that they may already have had a good ability of analysing social interactions and did not get much added value from the feedback in the thought bubbles. Alternatively, it may have been the case that the provided feedback was not explicit enough. The SGLM suggests that players need to be drawn out of the game world to reflect on a meta-level. Our implementation of thought bubbles in the game may not have achieved this effect, which calls for closer scrutiny of how effective feedback should be provided. It could also be that implementations of feedback should be tailored to player preferences to give them the support they would benefit most from.

Based on the analysis of the SJTs, we cannot prove Hypothesis 2 and 3 either. Thought bubbles did not improve the participants' understanding of how their behaviour invites behaviour of the virtual juvenile. Neither did LOITER-TB assist participants in recognizing stances of other people. Because the police officers were experienced at street interventions, they already were adept at recognising stances. Therefore, there may be a ceiling effect, as they also indicated that they did not learn very much ($M = 3.24, SD = 1.57$ on a 7-point Likert scale). We expect that the learning effect of LOITER-TB may be stronger for less experienced police officers or trainees.

Participants commented on the difficulty of rating the word pairs for the PFnet analysis. Combined with the fact that we were not able to measure differences between the two conditions, we conclude that PFnet analyses may not be entirely suitable to determine people's mental models about social awareness.

Lastly, it may be the case that the game of LOITER simply does not offer the right fidelity to bring the experience of social conflicts across. Its cartoon style may work well for some participants, especially with the current, straightforward interactions, but it may also hamper its effects because trainees cannot relate to the game's design. As the game is relatively short, it may also not bring across the experience of a heated debate or conflict. Furthermore, the game's design knows several limitations which may also adversely affect the gaming experience. For example, players can only input their choice by choosing from one of four predetermined choices, which may not necessarily match their actual intentions. Additionally, the virtual juvenile behaves in a way which may become predictable to some players, as he will always become friendlier as long as the player acts friendly. In the following section, I address the design of LOITER-FB,

Figure 34.: The second design of feedback in LOITER-FB.

which includes an implementation of the meta-technique of flashbacks as well as further improvements to LOITER-TB.

6.5 FLASHBACKS

LOITER-FB

In a second prototype called LOITER-FB (*FlashBack*), we implemented a new meta-technique, flashbacks, and addressed several shortcomings of LOITER-TB. We designed the flashbacks to mimic their use in larp: they refer to events that took place earlier in the time line of the game's story and reveal more about the character's cognitive state. In our implementation, the flashbacks provides feedback on the players' behaviour during the game. A flashback shows a past action and explains how the player's past average stance influences the juvenile's current preferred stance. Thus, the flashback provides both a retrospect on what happened and a prospect to what may happen. We implemented flashbacks as thoughts of the juvenile and presented them in thought bubbles, see Figure 34a. As in LOITER-TB, the overarching learning objective of LOITER-FB is improving the social awareness of players.

To enable the use of flashbacks, we wrote two additional scenarios for LOITER-FB: one about the juveniles bothering shopping public, and one about the juveniles loitering at a business park. Thus, players interact with the group three times and have to bond with the juveniles over time. This way, LOITER-FB aims to overcome the limitation of LOITER-TB of having only one short scenario in which players interact with the juveniles. Additionally, we modified the cognitive model of the virtual juvenile so that it takes the player's behaviour during previous scenarios in the same game into account. This modification aims to stimulate players to become aware that their actions have an effect on the longer term. The flashbacks are aimed at instilling this awareness by providing feedback that contains information about the effects of the player's actions on the virtual juvenile.

Internally, the player's choices correspond to stances in a two-dimensional grid resembling Leary's Rose (see Section 5.2.3). When a scenario has come to an end, the average of the player's choices is calculated as a point somewhere in that

grid. Based on this average, the juvenile will have a preferred stance for the next encounter, namely the one that is invited by this average (see Section 5.2.3). For example, the player's average stance could be $(2, 1)$, which would correspond to a very friendly, slightly dominant stance (as the axes in the grid range from -2.5 to 2.5). Based on this average, the juvenile's preferred stance would be $(2, -1)$, corresponding to a very friendly, slightly submissive stance. In the second and third scenario, the juvenile's response to the player's actions is calculated as an average of the stance that is invited by the player's action and the juvenile's preferred stance. For example, the juvenile's preferred stance could be $(2, -1)$, the player's action $(2.5, -2.5)$ and, following from this, the invited stance would be $(2.5, 2.5)$. The resulting new stance of the juvenile would be the average of the first and last coordinates: $(\frac{2+2.5}{2}, \frac{-1+2.5}{2}) = (2.25, .75)$. This would be a very friendly and slightly dominant stance, as opposed to the invited stance, which would be at the extremes of friendliness and dominance in the grid.

For LOITER-FB, we adapted several other parts in LOITER-TB's design. Based on comments from participants and police instructors who played LOITER-TB, we created an alternative visualization of Leary's Rose in the thought bubbles of the virtual juvenile, see Figure 34b. Instead of showing pictures of animals, the thought bubbles show Leary's Rose with an emphasis on the stance that the player assumed. To prevent information overload, we show the thought bubble and the juvenile's reaction in sequence, and not at the same time.

6.6 EXPERIMENT: FLASHBACKS IN LOITER-FB

To evaluate our implementation of flashbacks, as well as the additional modifications to the game's design, we conducted an experiment with LOITER-FB similar to that of LOITER-TB. We expected the flashbacks to assist trainees in improving their social awareness. Our hypotheses were the following, analogous to those of the evaluation of LOITER-TB.

HYPOTHESIS 1 Flashbacks help trainees to attain knowledge about social interaction that is similar to that of domain experts.

HYPOTHESIS 2 Flashbacks help trainees to determine which kinds of behaviour their own behaviour invites.

6.6.1 Method

For this experiment, we again used a repeated-measures, between-subjects design with two conditions. We conducted this experiment with police trainees in the second year of their curriculum. One group played the game with thought bubbles and flashbacks (condition FB) and the other played the game without (condition C). In total, participants made 24 choices (8 choices for each of the three scenarios) when interacting with the juvenile. In condition FB, 12 thought bubbles of the virtual juvenile were shown as well as two flashbacks at the beginning of the second and third scenarios.

Instead of using PFnet analyses, we investigated participants' learning experiences using two measures, both improved versions of the SJTs used for the experiment with LOITER-TB. Again, we created these SJTs in cooperation with an instructor of the Dutch Police Academy. Compared to the SJTs used in the previous experiment, we provided additional context in the situation descriptions and let participants rate the given approaches instead of selecting only the best option from them.

In the first SJT, participants rated the effectiveness of four approaches of a police officer (based on the four stances) when he confronts a group of loitering juveniles. The participants were given the description of a real-life scenario, followed by four approaches they were asked to rate. One of the two descriptions is shown in Transcript 18 (translated from Dutch, see Appendix B.2). Participants rated the effectiveness of four such approaches for each of two different scenario descriptions, for a total of 8 ratings on a Likert scale from 1 to 7 (measure SJT1). This measure was aimed at answering Hypothesis 1.

A group of teenagers is standing around a bench in a shopping street.

You are a patrolling officer, passing by a bit further down the street. You see the teenagers are staring at a woman passing them by. When she passes them, one of the teenagers whistles at her. The woman abruptly throws her head back, to which the teenager reacts: 'Hey, you feel like coming with us?' The woman hastens in your direction and ignores the group, after which the teenager calls after her: 'Slut!' The whole group cannot contain their laughter. The woman shakes her head while she exits the street: 'Bunch of idiots.'

- (A) You calmly approach the teenagers. You tell them you won't accept their cursing and you ask them how all of you can make sure this won't happen again. (DF)
- (B) You wait for a minute and then calmly approach the juveniles. You ask them whether they have been harassed previously and how that came across to them. (SF)
- (C) You whistle at the juveniles and shake your head. You walk by them and gesture that they should better go somewhere else. (SH)
- (D) You directly step towards the juveniles and tell them should behave normally. You don't tolerate the harassment of people. The choice is theirs: get out of here or get a fine. (DH)

Transcript 18.: The first scenario of SJT1 for LOITER-FB, showing the description of a setting and four different approaches of the police officer. The indications of the four different stances were not shown to the participants.

In the second SJT, participants rated how likely an utterance of a juvenile would be a response to an utterance of a police officer. First, participants read

Table 11.: Participants' scores on the situational judgement tests, per condition.

CONDITION	SJT1 SCORE, 8 ITEMS (M, SD)			SJT2 SCORE, 8 ITEMS (M, SD)		
	PRE	POST	INCR.	PRE	POST	INCR.
C	9.4 (2.3)	8.9 (2.7)	-.5 (3.1)	19.3 (4.9)	22.4 (4.4)	3.1 (4.6)
FB	9.3 (4.3)	9.7 (2.8)	.3 (4.7)	23.1 (5.9)	23.2 (6.2)	.1 (4.3)

through a short scenario, similar to the above one. Then, participants were given a description of the approach the police officer took, for example, a submissive-hostile stance: *'You keep on walking, shaking your head. You say, just loud enough for the teenagers to hear: "Well, today's youth certainly isn't what it used to be in my days..."'* Participants then rated the likeliness of four different behaviours of the juveniles on a Likert scale from 1 to 7. Appendix B.2 lists this SJT in its entirety, consisting of four police behaviours with four possible responses of juveniles for a total of 16 items (measure SJT2). The SJT2 measure was aimed at answering Hypothesis 2.

We compared the responses of participants in both SJTs to those of domain experts (two police instructors), rather than relying on preconceived notions about how these situations should be judged. The ratings of the domain experts did not differ significantly (t -test, $p > .05$). We constructed similarity scores based on the differences between the ratings of the participants and the average of the ratings of the domain experts. At the end of the experiment, participants filled in the same self-report questionnaire as in the experiment on LOITER-TB.

In total, 28 senior police trainees completed the experiment. Their mean age was 25.9 (SD = 7.3); 72% was male and 28% female. All participants completed the experiment within 30 minutes; 15 played condition c and 13 condition FB.

6.6.2 Results

Table 11 shows the means of the participants' similarity scores on both SJTs and their increase in scores.³ Independent samples t -tests did not show a difference between conditions on initial scores ($t(24) < 1, p > .05$), nor on post-game scores ($t(24) < 1, p > .05$) on SJT1. A one-sample t -test ($t(25) < 1, p > .05$) showed that the change in the scores of all participants did not significantly differ from 0. An ANCOVA (independent variable: post-SJT1 score; fixed factor: condition; covariate: pre-SJT1 score) showed no significant difference between condition c and FB ($F(25) < 1, p > .05$).

Independent samples t -tests did not show a difference between conditions on initial scores ($t(24) = -1.794, p > .05$), nor on post-game scores ($t(24) < 1, p > .05$) in SJT2. A one-sample t -test ($t(26) = 1.942, p > .05$) showed that the change in the scores of all participants did not significantly differ from 0. An ANCOVA (independent variable: post-SJT2 score; fixed factor: condition; covariate: pre-SJT2 score) showed no significant difference between condition c and FB ($F(27) < 1, p > .05$). In both conditions, participants were satisfied with how they solved

³ We assume linearity of the Likert-scale items to average data across participants.

the game's scenario ($M = 5.5, SD = .8$), but were not convinced that the game improved their understanding of how their behaviour influenced others ($M = 3.8, SD = 1.6$).

6.6.3 Discussion

The situational judgement tests (SJT1 and SJT2) did not prove Hypotheses 1 and 2. We cannot conclude that interactions with LOITER-FB affected trainees' social awareness, either positively or negatively. We did not find significant changes in the participants' social awareness, neither across, nor between the conditions. Our situational judgement tests showed that, before playing the game, participants already deviated only slightly from the average of the domain experts' ratings on both SJTs. This indicates that the participants may not be able to improve their social awareness much more, as our results confirm. Because of this ceiling effect, the added value of LOITER-FB may not be apparent for this user group. This may also have been a result of the participants' experience, as they were all trainees with some experience in the field concerning addressing social conflicts. As in the evaluation of LOITER-TB, this may have caused the presence of the ceiling effect in scores.

Some of LOITER-TB's limitations remained in LOITER-FB. The limited means for input made it easy for participants to tell which response would be the preferred one. They commented that they believed the game presents realistic cases, but that the difference between the available actions was very clear. Future prototypes may therefore benefit from a game mechanic that would allow for less restricted input, for example, allowing players to combine different nonverbal and verbal behaviours, or free text input. While we addressed the limitation of the game being short by creating two additional scenarios for a total of three scenarios, this still constituted relatively short interactions. Together with the game's style, this may again have not come across as a representative simulation of social conflicts in real life. One limitation of the setup of the experiment was that we did not evaluate to what extent participants processed the feedback presented in the flashbacks. It may be that they glanced over the presented flashbacks and did not reflect on their meaning. However, as the experiment showed that there was a ceiling effect involved in the learning gains, whether or not they reflected on the flashbacks would probably not have affected their learning experience. Like LOITER-TB, LOITER-FB may be more suitable for less experienced trainees, for example, first-year students with no practical experience.

6.7 REFLECTIONS ON THIS CHAPTER

In this chapter, I addressed the third Research Question of this dissertation:

RQ3 How can serious games stimulate meta-thinking to increase their effectiveness?

To answer this question, I discussed how I was inspired by meta-techniques in larps, which are often used as feedback mechanisms to both enrich the experience

of their players and to let them reflect on their experiences. In a similar way, meta-techniques can be used to stimulate reflection in our serious games. Reflection aids the process of learning. In the context of improving social awareness, reflection can help people analyse their own behaviour and how it influences the behaviour of others. For serious games, several types of feedback can be used to stimulate players' reflection. In the previous chapter, I described LOITER, a prototype serious game revolving around a police officer trying to resolve a situation in which loitering juveniles are causing a nuisance. Players take on the role of the police officer and have to resolve the conflict in a way that they see fit by choosing actions in a turn-based fashion: the player chooses an action and the juvenile responds to it. The juvenile does so based on the theory of interpersonal reflexes, which describes people's interpersonal *stances*: their attitudes towards each other.

To investigate how the learning process of police trainees can be supported with feedback in LOITER, I designed two modifications of this game which incorporated the meta-techniques of thought bubbles and flashbacks, respectively: LOITER-TB and LOITER-FB. These meta-techniques offer additional information to players that is not directly available to them in LOITER. In LOITER-FB, the thought bubbles give insight into the virtual juvenile's thoughts about the player's approach. When a player has taken an action, the juvenile's thought bubble shows the *stance* of the player's action in terms of the theory on interpersonal reflexes. We hypothesized that such feedback would improve trainees' social awareness as it allows them to reflect on how their behaviour came across to the juvenile and affected his response. Similarly, the inclusion of flashbacks was intended to emphasize the causal links that exist between the player's actions and the juvenile's responses. The second modification, LOITER-FB, comprises three scenarios in which the player encounters the same group of loitering juveniles thrice. Flashbacks show how the player's interactions with the juveniles in a previous scenario are remembered by the juvenile's in the current scenario. We hypothesized that this information would help players to understand how their behaviour influenced that of the juvenile in the longer term.

While there seems slight evidence that these forms of feedback improve trainees' social awareness, we did not manage to find anything conclusive concerning the learning gains of participants for either implementation. We compared models of relations of concepts of participants to those of domain experts and found no significant differences between them; neither before nor after the participants had completed the game. The same held for comparisons of the answers given by participants and domain experts to situational judgement tests. We suspect that the main cause for this result is that participants were already experienced with social conflicts and could thus not benefit much from their interactions with LOITER. It may be that these implementations of feedback result in improved reflection of less experienced trainees, but investigating this is left for future research.

The research presented in this chapter opens up possibilities for more investigations into how various facets of LOITER's design can be improved. Firstly, we did not directly inspect how players interacted with the feedback presented in the game. It may be that they read it thoroughly, glanced over

it, or ignored it altogether. Unobtrusive means, such as eye trackers, can be used to see how much attention the feedback receives from players. Secondly, interaction with LOITER is limited to selecting actions from pre-defined sets; this may be improved by allowing for different input methods, such as (restricted) text input or verbal input. A large challenge in this approach is the automated recognition of player stances, which has previously shown to be a challenging task (Vaassen & Wauters, 2012). While the evaluations with LOITER-TB and LOITER-FB did not result in directly measurable results, these implementations of feedback techniques employ a paradigm worthy of further exploration. Thirdly, the meta-techniques used in LOITER provide additional information to players; as with all feedback, the crux is in optimizing the content and timing of these techniques. The current implementations of thought bubbles and flashbacks are presented in a way that is intended to not intrude upon the gaming experience, while still providing players with the possibility to reflect on their experience. The degree to which feedback intrudes upon the gaming experience can depend on a player's competence; more experienced players will not benefit that much from feedback, while less experienced players might urgently need it. Tailoring the timing and content of feedback to individual players may therefore help to make the game a more effective learning experience. Fourthly, player reflection is only stimulated via implicit means. Instead, game mechanics can be designed to enforce explicit reflection; as discussed previously, this entails a decrease in immersion (cf. the Serious Games Lemniscate Model (Koops & Hoevenaars, 2012)). Ongoing work should investigate to what extent such intrusive mechanics influence the experience of players, both in terms of learning and of engagement, in search of the optimal division between the two.

In the following chapter, I investigate a related form of tailoring serious games to players to create improve the efficacy of LOITER. I address how the difficulty of serious games can be adapted dynamically to tailor the learning experience to individual players.

What can change the nature of a man?

Ravel
Planescape: Torment
 Black Isle Studios (1999)

IN this chapter, I discuss how the meta matters to virtual agents in serious games.¹ This chapter revolves around Research Question 4 of this dissertation:

RQ4 *How can virtual agents use meta-thinking to adapt serious games to the user's competence?*

To answer this research question, I discuss how serious games can dynamically adapt to suit learners' educational needs, showcasing an evaluated implementation in LOITER. As learners may learn in different ways at different speeds, serious games can benefit from adapting their contents to a learner's style of learning and his or her competence (Bellotti, Berta, De Gloria & Primavera, 2009; Tseng, Chu, Hwang & Tsai, 2008). I limit my investigation into adaptation in serious games to the adaptation of the game's challenge to the player's competence. First, in Section 7.1, I discuss the theories formalizing this relation and previous work on tailoring challenge to competence. In Section 7.2, I describe which parts of LOITER are suitable for adaptation. To provide an appropriate challenge to players, I propose the use of *adaptive profiles* for virtual characters in a third LOITER prototype called LOITER-AP. Adaptive profiles enable virtual characters to behave less cooperatively, which makes it harder for more advanced players to convince the characters to cooperate. To adapt their profiles, the virtual characters go meta by determining whether players are competent enough to progress to a higher level of difficulty. This way, LOITER-AP becomes more difficult when players are more competent. In Section 7.3, I describe the evaluation of LOITER-AP. Results of this evaluation show that, while the use of adaptive profiles does not increase the game's challenge, players felt improved behavioural involvement with the virtual characters. I close this chapter by discussing how the meta matters when agents adapt their personalities, the limitations of my approach and the possibilities for future research in Section 7.4.

ADAPTIVE
 PROFILE

7

¹ This chapter is inspired by Linssen and Theune (2012).

7.1 ADAPTIVE DIFFICULTY IN SERIOUS GAMES

FLOW

Games provide challenges to players that they need to overcome in order to reach objectives in the game, for example, to bring a story to a happy ending. Through playing, players learn and improve their skills so that they are able to overcome increasingly difficult challenges. This is one of the key elements that makes games entertaining: they provide challenges that match a player's current skills. As I explained in Chapter 5, games are objects that exemplify Csíkszentmihályi's thoughts about *flow* they provide specific challenges and enable players to lose their sense of the outside world (Csíkszentmihályi, 1990). In this section, I explain how a key part of Csíkszentmihályi's flow theory, the matching of a person's competence with a task's challenge, has been investigated in the context of serious and commercial-of-the-shelf (COTS) games. First, I address the theory behind this balance between competence and challenge, and illustrate it using examples from COTS games. Second, I provide an overview of systems which adaptively provide players with different challenges.

7.1.1 Tailoring challenge to competence

In the context of serious games, Csíkszentmihályi's flow theory is often used to illustrate that tasks can be best performed when they suit a person's competencies (Bailey & Katchabaw, 2005; Breuer & Bente, 2010; Kickmeier-Rust, Mattheiss, Steiner & Albert, 2011). That is, players enjoy a game the most when the game's difficulty matches their skills. However, this is a limited view of Csíkszentmihályi's work. He intended this relation between competence and challenge to be used as an integral part of his preconditions for achieving a flow state. First and foremost, for people to be 'in the flow', they need clear objectives as well as insight into their progress. Secondly, people need comprehensible feedback on their progress, based on which they can determine whether they are on the right track. Finally, the balance between a person's skill level and the challenge of the task at hand determines whether that person can attain a flow state. Figure 35 shows the various mental states which, according to Csíkszentmihályi, may arise from the different combinations of a person's skill level and a task's challenge level. Precisely when a person has a high skill level and faces a highly challenging task, then she may enter the flow state.

ZONE OF
PROXIMAL DE-
VELOPMENT
(ZPD)

Highly related to this work is that of Vygotsky (1978b), who researched developmental psychology, specifically in the context of children's education. He postulated that, in order to achieve optimal development, there should be a proper balance between instruction and development. He captured this idea in his concept of the *Zone of Proximal Development* (ZPD). Vygotsky (1978a) defined the ZPD as the area between the tasks that a learner can do unaided and the tasks a learner cannot do at all. So, tasks in the ZPD are those that the learner can do with assistance. Because of his interest in the ongoing development of children, the ZPD has also been used to study the dynamic aspects of learning. Figure 36 gives a schematic interpretation of how the ZPD may change, based on a person's competence and the challenge of a task. This interpretation shows an overlap with the relation between

Figure 35.: Possible mental states for varying degrees of skill and challenge (adapted from Csíkszentmihályi (1990)).

these two variables as Csíkszentmihályi described. Like flow, people enter the ZPD when there is a good balance between competence and challenge.

With proper assistance, a learner remains in the ZPD to improve his skills, opening the way for bigger challenges. Vygotsky (1978a) asserts that such assistance is vital to keep learners pushing their limits and learning most optimally. The research into *scaffolding* builds heavily on this idea: learners need guidance and support from tutors (Stone, 1998). For example, tutors can simplify a problem or highlight key components of a problem solution (Wood, Bruner & Ross, 1976). Over time, this scaffolded guidance can be decreased to let learners perform independently. With the rise of computers, scaffolding can be automated in various ways (Quintana et al., 2009), for example, by providing hints on how to navigate a virtual environment (Laffey, Schmidt, Galyen & Stichter, 2012), feedback on emotional states of virtual characters (Gupta, Walker & Romano, 2007), and through the use of narratives (Riedl & Young, 2014).

SCAFFOLDING

Similarly, computer games lend themselves well to the automatic adaptation of game settings to players, although only a small portion actually implements adaptation to match their challenge to the player's competence. All digital games follow an action-reaction cycle in which the game world reacts to the player's actions. Yet relatively few games go meta and step outside of this cycle to determine how to respond to players. Here, I distinguish between three types of games when discussing adaptation: those that lack adaptation, those with interactive stories and those that take into account meta-information about the player's experience.

Figure 36.: A schematic interpretation of Vygotsky's Zone of Proximal Development (Vygotsky, 1978a).

Firstly, most games feature static mechanics that do not adapt to players' actions. For example, *DOOM*, a classical first-person shooter, features many levels which players have to traverse, filled with enemies they have to defeat (id Software, 1993). The levels and the enemies and items they contain are always the same, independent of previous actions of players. Enemies respond differently only in direct reaction to players' actions.

Secondly, as I explained in Chapter 2, some games, like those of the *Mass Effect* series (BioWare, 2012), do feature mechanics that let players influence how the story and, in turn, the challenges, play out. For example, based on the player's decisions in this series, characters may live or die, and either accompany the player in later instalments of the series or not. Such games carry out adaptations based on reasons within the story world: in-character motivations.

Finally, a third type of games steps out of the bounds of the magic circle and reasons about players' behaviour on a meta-level. For example, the shooter *Max Payne* (Remedy Entertainment, 2001) keeps track of several statistics of the player character, such as health and accuracy. Based on these variables, this game adapts the difficulty of the game, letting more enemies attack the player when she is up for the challenge or making them fire less accurately when she is having more trouble completing a level. Another example is that of *Mario Kart* (Nintendo, 1992), a racing game which features the well-known *rubber banding* mechanic (Pagulayan, Keeker, Wixon, Romero & Fuller, 2002). This mechanic limits the degree to which enemy, non-player racers can outrun or keep up with players, effectively letting them slow down or speed up as if tied to the player with a rubber band.

Linssen (2011, pp. 9–10) discusses more examples of such *dynamic difficulty adjustment* (DDA) in COTS games, based on meta-information. Through DDA, the difficulty of games can be kept in check, to ensure players' tasks have just the right amount of challenge. However, this does not always have its intended effect. A much-discussed example of DDA gone awry is that of *Oblivion*, a role-playing game which features a large open world in which players are tasked with a large variety of quests (Bethesda Game Studios, 2006). In *Oblivion*, players can improve the skills of their characters by completing these quests, becoming more powerful over time by levelling up. This game uses a scaling system which scales the properties of enemies to the level of the player, affecting all the enemies in the game. While this means that this task is just as challenging to players of any level, it also has the unwanted consequence that the sense of achievement gained from improving the player character's skills is nullified (Adams, 2008; McClure, 2009). This example shows that tailoring all challenges to players' competence may not always work out as intended.

In the following section, I address several investigations into the adaptation of the difficulty of serious games.

7.1.2 Implementations of adaptive difficulty in serious games

The difficulty of games can be adapted in different ways. In this section, I first describe how COTS games can be adapted before discussing adaptive difficulty in serious games.

Bailey and Katchabaw (2005) describe how several factors of a game can be adapted to tailor its difficulty to players. Bailey and Katchabaw distinguish between attributes of both player and non-player characters, such as their strength and health, or their thought process, and attributes of the environment and puzzles, such as the placement of items. Bailey and Katchabaw investigated how to dynamically adapt the difficulty of a shooter game, based on *Unreal* (Epic MegaGames & Digital Extremes, 1998). Their implementation of DDA allowed the game to support players by adding items to the environment, for example, ammunition for guns when players are short on ammo or health kits when players are low in health. They evaluated their implementation in a preliminary study through an experimental testbed which required players to overcome jumping, shooting and dodging challenges. One of their main conclusions was that their implementation of adaptation was solely reactive: it required players to commit a mistake or not complete a challenge in time. Bailey and Katchabaw emphasise that good adaptivity should also be proactive to guard players from having to repeat their task to complete it successfully.

A game's difficulty can also be adjusted through a scaffolding approach by providing players with explicit hints about a task. Peirce, Conlan and Wade (2008) investigated DDA through the *Adaptive Learning In Games through Non-invasion* system (ALIGN). They describe a generic approach to generating adaptations in various forms, such as providing different learning situations and providing support within learning tasks, which Peirce et al. call macro- and micro-adaptivity, respectively. To implement ALIGN, Peirce et al. used the ELEKTRA environment,

which was a result from a project which dealt with teaching concepts relating to physics. This happens through a first-person 3D environment which featured Galileo Galilei as a mentor (Kickmeier-Rust et al., 2006). Due to project constraints, Peirce et al. were only able to explore the use of micro-adaptivity in ELEKTRA, through motivational hinting and support and meta-cognitive feedback. The virtual embodiment of Galilei functioned as a tutor who expressed this support and feedback to the user of the ELEKTRA environment, an approach similar to *intelligent tutoring systems* (van Lehn, 2011). In this implementation, the adaptation relied on ALIGN detecting whether players had difficulty proceeding with a task and also detecting with which part of the solution they had trouble. For example, hinting at a solution in a particular setup, Galilei would point out that ‘Wood is a medium-weight and non-magnetic material. It is not influenced by the magnet, but can be influenced by the fan,’ (Peirce et al., 2008, p. 33). A small study revealed that users who were exposed to adaptivity improved their knowledge about the physics concepts treated in the virtual environment and were more confident in their own achievements.

Kickmeier-Rust, Steiner and Albert (2011) expanded the work on ALIGN by investigating macro-adaptivity in *80days* (Göbel, Mehm, Radke & Steinmetz, 2009). The virtual environment presented in *80days* features a friendly alien who wants to explore earth and could use some help in getting around. Players of this game learn about the geography of earth in collaboration with the alien, completing several missions before finishing the game. For this game, Kickmeier-Rust et al. implemented macro-adaptivity by adapting the order in which missions are presented to players, as well as providing micro-adaptivity through hints. The key factor here is that missions are presented to players when they align with (1) the players’ educational needs, which are determined through outcomes of earlier missions, and (2) the story thusfar, making sure that, for example, locations are not visited again when they have been visited, to keep the story flowing. This way, the storyline of the game is tailored to players to provide an optimal learning experience. To formalise their method, Kickmeier-Rust et al. defined *narrative learning objects*, which consist of narrative components and educational components, with prerequisites for insertion into the story, and motivational and collaborative components (for a complete description, see Kickmeier-Rust, Steiner and Albert (2011, p. 233)).

Bosse, Gerritsen, de Man and Tolmeijer (2015) use a similar approach to tailoring a game’s challenge to the skill level of players as part of the STRESS project (see Section 5.1). This project investigates the creation of a virtual reality environment with the goal of training public transport personnel to interact with aggressive passengers who, for example, insult the tram driver because of a delay. Users can de-escalate situations by choosing, in a multiple-choice fashion, the correct communication techniques, which depend on the type of aggression of the assailant. Therefore, users need learn how to distinguish between different types of aggression before choosing a course of action, see Bosse and Provoost (2015). For the adaptation of the difficulty of the cases presented to trainees, Bosse et al. change the possible answers from which the trainees can choose, namely different combinations of exemplary, acceptable, or unacceptable options. As in *80days*,

these cases are presented separately, based on players' earlier interactions. With higher levels of difficulty, the number of unacceptable options is increased and the acceptable actions cease to sufficiently convince the virtual agent to calm his manners. An implementation of this concept also allowed trainees to drop to a lower difficulty and get some more experience with certain cases before continuing to the next level.

Where Bosse et al. presented separate encounters to players, *Code Red: Triage* (CRT) takes a different approach. CRT is a serious game about the bombing of a railway station (van der Spek, 2011). The goal of the player is to find her way around the station and to correctly *triage* 18 victims of the bombing through a standard procedure, checking for vital signs which indicate different levels of injuries, entailing different urgencies for different casualties. The victims of the bombing are spread throughout the station that players need to traverse, which means that all the separate cases are present in the same scenario. When players interact with a victim, they are required to choose from eight options to check for signs of injuries. Depending on their choices in triaging, the game rates the players on a scale of 0 to 100, subtracting points for incorrect or wrongly ordered actions, and adding points for correct actions. To provide an adaptive experience to players of CRT, van Oostendorp, van der Spek and Linssen (2014) developed an implementation of dynamic difficulty adjustment to provide macro-adaptivity, but on a smaller scale than that of *80days* and *STRESS*. Each type of injury requires a specific order of actions to complete the triage, leading to six different complexities of triages. Each complexity level has a threshold score players need to exceed; attaining this score indicates that the player has learned enough about that level of complexity. When the player exceeds this threshold score for a victim in a certain complexity level, the game increments the difficulty level of the player's next case. This way, players only need to triage the cases from which they can gain additional knowledge, effectively removing the types of cases that they have already mastered. An evaluation of this approach showed that, on average, the time players spent on their task was reduced by approximately 50%, increasing the efficiency of the game while still retaining equal learning gains.

For *in-company emergency assistance* (ICEM) training, Peeters, van den Bosch, Meyer and Neerincx (2014) developed an approach to dynamically adapt the scenario of a training case to control its storyline and effectiveness. This exploratory research was carried out with live actors to determine its viability for virtual environments. The crux of this approach lies in the actors being able to switch between two versions of the same scenario, either being supportive to players or challenging them. A director had an unobtrusive overview of the training situation, and directed the actors without a trainee noticing it. Through discussion with domain experts, Peeters et al. identified explicit behavioural cues, based on which the director was instructed to adapt the situation, either increasing or decreasing its difficulty. Experienced instructors of ICEM analysed recorded fragments of interactions, concluding that directed fragments received a higher learning value than non-directed fragments.

The above systems take various approaches to teaching players about their particular domains. All but the last of these systems adapt their difficulty either by

CODE RED:
TRIAGE (CRT)

IN-COMPANY
EMERGENCY
ASSISTANCE
(ICEM)

presenting cases or scenarios in different orders or by providing hints. The work by Peeters et al. (2014) dynamically adapts the behaviour of a character itself while a scenario is being played. For the new prototype in the LOITER series, we take a similar approach to adapting the game's difficulty. Instead of selecting scenarios that match a player's educational needs or generating such scenarios on-the-fly, we go one step beyond these approaches and investigate how we can change characters' behaviour during scenarios. This has two main advantages over the alternative solutions. Firstly, scenarios in LOITER take a few minutes to complete; by adapting the character's behaviour during a scenario, the difficulty can be increased quicker. This way, the game will be more efficient than when adapting the difficulty after a scenario has been completed. Secondly, Chapter 6 showed that the provision of hints through thought bubbles in LOITER did not yield significant learning gains. Therefore, the adaptation of these hints may not necessarily alter the game's difficulty, though they may assist in keeping the interaction with the virtual character coherent. In the following section, I explain the chosen approach in more detail.

7.2 ADAPTIVE STANCE PROFILES OF VIRTUAL CHARACTERS

LOITER-AP I discuss the design, implementation and evaluation of LOITER-AP, an adaptive version of LOITER. This prototype game adapts to players' skills by having a virtual character adopt different personalities (Sections 7.2.1 and 7.2.2).

LOITER-AP has two learning objectives in addition to those of LOITER-TB and LOITER-FB. The first of these states that people do not necessarily behave according to the rules of a psychological model. As Leary (1957) explained, certain behaviour does not always lead to an equally certain response.

LEARNING OBJECTIVE 1 Become aware of the fact that Leary's Rose is only a model; people may behave differently from what the model predicts.

To make the player aware of the reasons why the character changed his behaviour, the character's thoughts are shown to the player. This is intended to maintain the coherence of the character's actions in the eyes of the player (Section 7.2.2.3). In Section 7.2.4, I describe the implementation of this concept.

In contrast to the previous LOITER games, the juvenile can take initiative in adopting a stance instead of merely following a player's lead. With his sudden change in stance, the juvenile may provoke a reaction of the player as Leary's Rose predicts, mirroring the juvenile's stance. This may lead to players responding with hostile reactions when the juvenile adopts a hostile stance, which is in conflict with the second learning objective of LOITER-AP:

LEARNING OBJECTIVE 2 Persevere in being friendly; do not let yourself be provoked to become hostile.

In order for players to continue performing well in the game, they should persevere in their friendly approach. Only then will they be able to progress to more difficult interactions, indicating that their competence has improved.

7.2.1 *Design considerations for adaptation*

Interactions in LOITER-TB and LOITER-FB are very straightforward, because the virtual juvenile responds in a deterministic way, based on Leary's theory of interpersonal attitudes. The evaluations from Chapter 6 showed that it was relatively easy for players to complete the interaction in a good way, due to the simplicity of the scenario, the participants' level of experience and the game's design. This lack of challenge was one of the possible causes for the limited educational effects of both LOITER games. The goal of LOITER-AP is to provide a game of which the challenge is tailored to the players' needs, following Vygotsky's idea of the ZPD (see Figure 36).

First, I address which aspects of LOITER can be modified to make its difficulty adapt to players. In previous experiments, participants indicated that the game was very straightforward: the virtual juvenile responded very predictably and the optimal choice of action was clear. When the participants acted friendly (SF or DF), the juvenile would always reciprocate (DF or SF, respectively). This, combined with both the initially neutral stance of the juvenile and the fact that most participants chose friendly actions, resulted in some participants never being exposed to unfriendly behaviour by the virtual juvenile. These limitations of the earlier design of LOITER suggested different ways to change the game's difficulty:

1. The game could be made less predictable, for example by providing the virtual juvenile with a limited degree of *agency* (Mason, 2013). That is, the juvenile gets the means to take initiative in choosing which stance to adopt, instead of acting purely reactively to players' actions. This seemingly contradicts the theory of Leary's interpersonal circumplex, yet that theory is only meant as an indication which behaviour people might adopt. Other factors may also influence the adoption of a different stance, for example, social roles and the presence of rapport.²
2. Players could input their choice of action in a different way. The findings in Chapter 6 indicated that players may have seen through the game mechanics of LOITER and simply selected actions with an image that shows a friendly police officer. The forms of adaptation proposed below can obfuscate which action is the most appropriate.
 - a) The presentation of the different choices for the player could be altered, for example by omitting the pictures of the accompanying nonverbal postures. This may however conceal the intended stances of the available actions, as the verbal utterances by themselves are ambiguous, making it unclear for players how the game would interpret their chosen actions.
 - b) The game could allow for free text input, letting players type in their responses without presenting them with limited options to choose from, some of which may be obviously 'right' or 'wrong'. This would open the game up for players, letting them choose exactly how

² See Section 5.2.1 and Bruijnes, Linssen et al. (2015).

to phrase and express their intentions instead of selecting actions that most closely resemble their intentions. This calls for automatic interpretation of players' input to correctly determine their intended stance. As shown in the research of Vaassen and Wauters (2012), this is a challenging task, as people's intentions can be expressed in ambiguous statements.

3. The game could alter the presentation of feedback. In both LOITER-TB and LOITER-FB, the feedback that was presented to players did not increase learning gains. For LOITER-AP, it makes sense to only provide feedback when it is useful for the player, which is in line with the theory of the ZPD: only provide scaffolding when needed, that is, when the task would be too difficult otherwise.

The main learning objective of the game remains the same: players need to gain insight into behaviour during interpersonal interactions. The first of the above possibilities for modification of LOITER offers more variety in the behaviour of the virtual character, making it a suitable candidate for adapting the game's difficulty. The second offers more variety for the players' behaviour, but, as mentioned, players' stances cannot be sufficiently reliably determined from free text input. The last of the possibilities involves changing the amount of scaffolding the game provides. Yet, as became clear in Chapter 6, the forms of feedback implemented in LOITER-TB and LOITER-FB did not lead to significant differences in learning effects for players. This may have been a result of the games already being easy enough to be completed 'perfectly'. Therefore, it is unlikely that varying the scaffolding in the new prototype will make it more effective. As the adaptation of the behaviour of the virtual characters appears the most promising, I explore its implications for the design of LOITER-AP.

7.2.2 Final design of adaptation

I introduce the use of a *profile* in our game, defined as the preferred stance of a character, classified as a point in Leary's Rose (Gurtman, 2009). To modify the difficulty, we can use a profile for the virtual juvenile as a sort of mood or inclination: the juvenile's reaction is based on a combination of the stance that is invited by the player's action and the juvenile's profile. This lines up with Leary's theory about interpersonal profiles (Leary, 1957). As explained above, this means that the juvenile's stance does not move in the coordinate system of Leary's Rose with a fixed distance, as was the case in previous LOITER games. Instead, the juvenile's profile influences how much his stance changes. For example, a juvenile with a dominant-hostile profile will respond more intensely to a player's hostile behaviour. This means that this juvenile would more rapidly be invited to behave dominant-hostile as compared to a juvenile with a neutral profile, with no preference for any stance (as in previous versions of LOITER).

The design of the adaptation comprises three main parts: the influence of the stance profile on the character's actions, the calculation of the profile, and the timing of the adaptation.

7.2.2.1 The stance profile

The juvenile's profile consists of two properties which influence his cognitive process of determining a new stance. The first is his *stance profile*, classified as a point p_s in Leary's Rose; the second is a factor p that determines to which extent the juvenile prefers that stance. The profile value p is a number equal to or larger than 1, which either lengthens or shortens the vector that express a change of stance in Leary's Rose, see Figure 37. The profile functions similarly to a rubber band: it pulls the invited stance to the juvenile's preferred stance. Without such a profile, or with a neutral profile with a profile point $p_s = (0,0)$ (the origin of the axes) and a p value of 1, Leary's Rose dictates the following behaviour. When the player takes a submissive-hostile action a_1 , he invites the juvenile to act dominant-hostile (i_1). With the neutral profile, the juvenile's current stance (s) would change with a vector s_1 in the direction of i_1 .

STANCE
PROFILE

When the juvenile has a preferred stance, for example, a point in Leary's Rose in the dominant-hostile quarter (p_s), the resulting stance will be different. The non-neutral profile p_s serves as an attracting force which influences the vector's length, either lengthening or shortening this vector, depending on the player's action and p . For the player's action a_1 , the p value is multiplied with the vector s_1 , because s_1 points in the direction of p_s , resulting in a vector s_{p1} . Through the use of this profile, the juvenile gets dominant-hostile more rapidly when he is invited to adopt that stance, compared to a juvenile with a neutral profile. Conversely, when the player takes an action a_2 , he invites the stance i_2 , which is submissive-friendly. This is the opposite of the juvenile's profile. The original vector s_2 is divided by p , leading to a vector s_{p2} which is shorter than the original. This way, the juvenile's profile withholds him from becoming submissive-friendly as fast as a juvenile with a neutral profile. When the juvenile is invited to adopt a submissive-hostile stance, the vector's friendliness dimension is multiplied with p and its dominance dimension is divided by p . When the player invites the juvenile to act dominant-friendly, the inverse holds. Appendix B.3 contains the pseudocode of the algorithm governing the calculation of the dominance of a new stance.

7.2.2.2 Calculation of the profile

In LOITER-TB, the virtual juvenile remembers its past interactions with the player and adopts the stance the player on average invited with her past actions. This means that, if a player performed well, the juvenile would be friendlier in the ensuing scenarios as well, with the opposite also holding true. This may have been counter-productive as it made the game easier for good players and more difficult for bad players. For LOITER-AP, the inverse needs to hold. A juvenile's profile needs to correspond to a level of difficulty that matches with the players' competence. The game should not become easier over time, but more difficult, if the player performs well. Players who achieve a friendly relationship with the virtual character can progress to the next level of difficulty, namely one in which the virtual character is less inclined to cooperate.

To increase the difficulty, the juvenile needs to adopt an unfriendly stance. This way, the player needs to put in more effort to make the juvenile adopt a friendly

Figure 37.: The effects of the p value in the calculation of a new stance.

stance again. While playing the game, the juvenile's stance is an indication of how well the player is performing: when he adopts a friendly stance, the player has been doing well. When the juvenile is friendly for a sufficient amount of time or at a certain point in time, LOITER-AP's difficulty can be incremented by letting the juvenile adopt an unfriendly profile. Conversely, the difficulty is not incremented when the juvenile is in an unfriendly stance, as in that case the player apparently was not up to the task of making him cooperate. We decided not to decrement the game's difficulty as previous studies showed that the game's basic difficulty already was low enough for most players to complete the game successfully (see Chapter 6). Table 13 shows how the juvenile's stance changes, depending on his current stance. Players' actions invite the juvenile to adopt a stance, yet the adaptation makes the juvenile adopt the inverse of the invited stance. For example, when a player acts submissive-friendly, the juvenile is invited to act dominant-friendly. However, the algorithm for adaptation makes the juvenile adopt a submissive-hostile stance: the opposite of a dominant-friendly stance. We base this on the notion that this is the maximally unexpected behaviour for the juvenile, which would take players by surprise. This is in line with the following learning objective 1 of LOITER-AP.

7.2.2.3 Timing and difficulty

Adaptation can take place at different moments during gameplay. With the current setup of LOITER, two types of adaptation are viable: within and between scenarios. The first provides (sudden) changes within a scenario, which we expect to be less expected than stance changes in between scenarios. The second lets players deal with juveniles who have stances different from neutral at the start of a scenario.

Players then will have to learn how to cope with juveniles with such profiles. In total, there are three scenarios players can play in LOITER-AP. We chose to include one adaptation within and one between each scenario, leading to five potential moments of adaptation. This entails a total of seven difficulty levels for LOITER-AP: those following the five moments of adaptation, and those at the start and end of the game.

To explain how the difficulty grows gradually over the course of the game, we first explain the design of the maximal level of difficulty. Adaptations take place exactly halfway a scenario: after 4 player actions in this scenario, after which 4 actions remain for the player. In this span, players should be able to get the juvenile in a friendly stance: the juvenile's affiliation should be positive (greater than 0). The juvenile's stance determines how many steps this would minimally take the player. For the highest difficulty in the game, the juvenile's affiliation will be -2.5 (LOITER's lower limit of both dimensions of the framework for the interpersonal circumplex). From this, it follows that the player needs to move the juvenile's position in Leary's Rose by more than 2.5 in four steps. A straightforward calculation shows that p should be smaller than $(4/2.5 \Rightarrow) 1.6$; we chose 1.5 as a readable number which covers this requirement. Table 12 shows the p values and stance for the juvenile's profile that correspond to the seven difficulty levels (cf. Figure 37). Here, the increase of difficulty relies on the minimum number of steps it takes players to get the virtual juvenile to adopt a friendly stance again. The different positions in the stance framework, combined with the different p values, allow for a gradual increase in difficulty. The difference between difficulty levels 2 and 3, and 4 and 5 lies in the fact that, in the third and fifth difficulty levels, players' actions will have less positive effects than in the second and fourth levels, respectively, because of the higher p value. That is, the juvenile's stance will move faster to the hostile end of the spectrum and slower to the friendly end. Therefore, when players take a hostile action, this will cause the juvenile to become hostile more rapidly in level 3 and 5 than in level 2 and 4, respectively). The choice between the pairs of coordinates the juvenile can adopt as a stance relies on the players' stance as explained in Table 13.

7.2.3 *Justifications*

The forms of adaptation described above may increase the difficulty of LOITER-AP, yet they may also be confusing to players. Suddenly, the juvenile does not immediately respond according to the preconceived rules (from Leary's Rose) anymore. Moreover, the reactions of the virtual character may seem counter-intuitive, especially when the player has been behaving friendly and the juvenile suddenly becomes hostile. To ensure perceived coherency of the juvenile's behaviour, the game can provide feedback on the reasoning behind the adaptation of his behaviour. Therefore, we include *justifications* for the juvenile's changes in behaviour to inform players about the causal links between their actions or other events and the juvenile's reactions. A similar approach is discussed by Szilas (2005), who notes that a character is believable when its behaviour is explainable and thus, that a character needs the means to express its justifications. As in LOITER-TB and

Table 12.: The seven difficulty levels in LOITER-AP, including the values for p and the stances adopted by the juvenile. Reaching the seventh level indicates that a player successfully completed the sixth level.

DIFFICULTY LEVEL	P	STANCE COORDINATES OF THE JUVENILE		MIN. NUMBER OF STEPS TOWARD FRIENDLY STANCE
		DH	SH	
1	1	(0, 0)	(0, 0)	1
2	1	(−1.25, 1.25)	(−1.25, −1.25)	2
3	1.2	(−1.25, 1.25)	(−1.25, −1.25)	2
4	1.2	(−2.5, 2.5)	(−2.5, −2.5)	3
5	1.5	(−1.25, 1.25)	(−1.25, −1.25)	3
6	1.5	(−2.5, −2.5)	(−2.5, 2.5)	4
7	–	–	–	–

Table 13.: The adaptation of the virtual juvenile’s stance, based on his invited stance and the player’s stance which invited that stance.

PLAYER’S STANCE	JUVENILE’S STANCE	INVITED	JUVENILE’S ADOPTED STANCE PROFILE
SF	DF		SH
DF	SF		DH
DH	SH		SH
SH	DH		DH

LOITER-FB, we choose the meta-techniques of thought bubbles and flashbacks to convey these justifications to players.

The justifications can refer to events internal or external to the conversation. This method builds on the out-of-character (OOC) approach discussed in Chapter 4. In the VST, characters reasoned that they should intentionally do something 'wrong' in order to make stories more captivating. In LOITER-AP, the juvenile can misinterpret things the player says or perceive a friendly attitude as being 'too much'. This misinterpretation is intentional from an OOC point of view, yet serves as an in-character explanation of the juvenile's behaviour to keep his behaviour coherent in the eyes of the player. Additionally, the juvenile can think about events outside of the conversation to keep his behaviour explainable. For example, his phone might be broken or he may have had a dispute with family members.

As explained in the previous section, adaptations can occur once during each of the three scenarios and two times in between scenarios. For LOITER-AP, we use thought bubbles for justifications of within-scenario adaptation and flashbacks for justifications of between-scenario adaptation. The justifications should explain the change in the juvenile's (preferred) stance. They do so through expressing (1) which stance the player adopted and (2) which stance the juvenile adopts (both as a profile and as a new stance). This holds for both types of adaptations, within and between scenarios. The content of the utterances is based on the juvenile's needs to 'save his face' (inspired by Goffman (1959) and Smelt, Habchi, Dekker and van der Wal (2014), see Section 5.2.1). The juvenile does not want to feel limited in his autonomy, nor does he want to feel ridiculed by the police officer, especially in the company of his friends. Therefore, the justifications will express his need to save face.

Each within-scenario justification is a compound of two sentences, the first expressing the juvenile's (mis)interpretation of the player's stance, the second expressing the juvenile's responding stance, indicating his face needs. Both sentences depend on the juvenile's responding stance. For example, when the player is acting DF, the juvenile will start acting as DH (see Table 13). This calls for a dominant-hostile misinterpretation of the dominant-friendly stance, for example, 'Suuure, this guy is gonna tell us what to do.' The second sentence follows up on this interpretation by conveying the DH stance, for example, 'Piss off, popo!' For the adaptations in LOITER-AP, all the interpretations are negative, as the juvenile only adapts his stance profile to a hostile stance (either DH or SH).

We created a list of example justifications which three police instructors of the DPA rated on the dimensions of dominance and friendliness of the sentences. Based on these ratings, we chose the six justifications that were rated as corresponding best to a particular stance (DH or SH) and fit the scenarios best, see Table 14.

The construction of the between-scenario justifications followed a similar process. These justifications also consist of a compound of two sentences, expressing both the juvenile's (mis)interpretation of the player's stance and the juvenile's responding stance. The key difference between between- and within-scenario justifications is that the former also refer to events outside of the context of the in-game interaction. These events are intended to explain the juvenile's interpretation of the player's course of action. For example, one of the events

Table 14.: The six different justifications that can be shown in thought bubbles when a within-scenario adaptation is made (translated from Dutch, see Appendix B.3).

SCENARIO	INTENDED JUVENILE'S STANCE	PLAYER'S STANCE	JUSTIFICATION
1	DH	DF	'Right, this guy is gonna tell us what to do. Piss off, popo!'
	SH	SF	'Damn, such a wuss. Don't pay attention to that loser.'
2	DH	DF	'Is this guy looking for trouble? He shouldn't fuck with me, man!'
	SH	SF	'You gotta be kidding, mister nice guy is here to bother us. Fuck this shit.'
3	DH	DF	'Such a commander, we're not in the army! If he's looking for trouble, come at me!'
	SH	SF	'Shit, he's feeling like it... Damn, if only he left us alone.'

the juvenile may refer to in a between-scenario justification is that of his parents seizing his phone. As explained above, the juvenile perceives this as a threat to his face need for freedom. The juvenile may then transfer his negative opinion about his parents to the player, who might currently also attack his freedom. In LOITER-AP, these justifications are shown as flashbacks at the start of the second and third scenario. They indicate the initial mood of the virtual juvenile for that scenario if the difficulty of the game has been incremented. As above, we created several justifications intended to convey the stances that correspond to the profile the juvenile adopts, see Table 15.

7.2.4 Implementation of LOITER-AP

In Section 5.2.3, I explained the design of LOITER, on which I expanded in Sections 6.3 and 6.5 with the descriptions of thought bubbles and flashbacks, respectively. As explained above, LOITER-AP uses the same basic setup. When the game starts, players get an introduction to both the goal of the game and its mechanics, see Figure 38. This is followed by an introduction of Scenario 1, which players subsequently play. At this point, the virtual juvenile has a neutral profile, as described in Section 7.2.2.3. After the user has chosen four actions (halfway the scenario, indicated by a grey, dashed line, see Figure 39), the game adapts its difficulty based on the stance of the juvenile. If he is in a friendly stance (either DF or SF), the difficulty is incremented; otherwise, the difficulty remains on the same level. Only if the difficulty is incremented, the game displays a thought bubble (TB) which provides a justification for the juvenile's change in stance, see Figure 39a.

Table 15.: The four different justifications that can be shown in flashbacks when a between-scenario adaptation is made (translated from Dutch, see Appendix B.3.2).

SCENARIO	INTENDED STANCE	JUSTIFICATION
2	DH	‘This morning I fucked up everything in the supermarket, now this guy is gonna whine as well? They should just leave me alone for once!’
	SH	‘Fuck, this morning my manager was getting on my back while I was stacking the shelves, now Mr. Cop. What a fucking mess...’
3	DH	‘Lost my phone to my parents, now this guy is gonna screw me? They should quit fucking me over!’
	SH	‘Oh shit, this keeps getting better. First my phone gets nicked by my parents, now this guy. Fuck, I don’t care for this at all...’

Figure 38.: A schematic overview of the different stages of a LOITER-AP playthrough.

The player can then continue choosing another four actions to take, after which the scenario ends and the game displays feedback on the stance both the juvenile and the player took on average. For the second scenario, another introduction is shown, after which a new scene is presented in the game. If the stance the juvenile assumed at the end of the previous scenario was friendly, the game increments the difficulty and shows a flashback (FB) at the start of the second scenario, see Figure 39b. As in the first scenario, an adaptation can take place after four actions, and at the end of this scenario, based on players’ actions. The same goes for the final scenario. The outcome of that scenario does not entail an adaptation, but serves as an indication whether the player successfully mastered the final and maximal difficulty level.

7.3 EXPERIMENT: ADAPTATION IN LOITER-AP

To evaluate LOITER-AP, I conducted an experiment to investigate the hypotheses underlying the game’s design. I expected players to experience an increase in difficulty when the juvenile’s profile was adapted. Players had to take more positive

(a) A thought bubble halfway through a scenario.

(b) A flashback at the start of a scenario.

Figure 39.: Two screenshots of the different types of justifications in LOITER-AP.

actions to get him to collaborate, so it was harder to get the juvenile in a positive, collaborative stance. This yielded the following two hypotheses.

HYPOTHESIS 1 Adapting the juvenile's stance (through stance profiles) will increase the objective difficulty of LOITER-AP.

HYPOTHESIS 2 Adapting the juvenile's stance (through stance profiles) will increase the subjective difficulty of LOITER-AP.

Furthermore, players might be invited by the juvenile's actions to mirror his stance. For example, they might get frustrated with the juvenile when he is being dominant-hostile. This yields the following hypothesis.

HYPOTHESIS 3 Adapting the juvenile's stance profile will invite players to adopt complementary stances, according to Leary's theory.

The juvenile's justifications were intended to keep the juvenile's behaviour believable despite changes of stance which do not follow Leary's theory. The thought bubbles and flashbacks were meant to make clear to players why the juvenile took a certain course of action. I expected these justifications to make the juvenile's behaviour seem more coherent to the player. While he might be seen as inconsistent without these justifications, he should now be seen as acting in a coherent way with causal relations between the player's actions and the juvenile's thoughts and reactions. This yields the following hypothesis.

HYPOTHESIS 4 Characters will be perceived as believable when justifications are provided for changes in stance that are not in accordance with Leary's Rose.

7.3.1 Method

This experiment is similar to the experiments described in Chapter 6: it uses a between-subjects design, this time with three conditions, corresponding to different versions of LOITER-AP. The first condition uses the implementation of

Table 16.: The structure of the Game Experience Questionnaire.

MODULE	COMPONENTS	# ITEMS	EXAMPLE ITEM
GEQ-core	Annoyance, Challenge, competence, flow, negative affect, positive affect, sensory and imaginative immersion	33	'I was fully occupied with the game.'
GEQ-ig	Same as GEQ-core, with a subset of its items	14	'I felt successful.'
GEQ-sp	Behavioural involvement, Psychological involvement: empathy, psychological involvement: negative feelings	17	'My actions depended on the other's actions.'

adaptivity as described above, including the justifications (condition APJ). The second condition includes the adaptive difficulty, but lacks the justifications (condition AP). The control condition (c) lacks both implementations and functions as a baseline.

The outline for this experiment closely mimics that of the experiments described in Chapter 6. First, participants received explanations about the general objective of this experiment, namely evaluating a method meant for training their interpersonal skills. They were then asked to fill out a consent form and start playing LOITER-AP in a web browser. The game itself included instructions on how to play it, including explanations about the justifications for condition APJ, detailing when the meta-techniques of thought bubbles and flashbacks were used and what they meant (see Appendix B.3).

We used several measures to test the hypotheses. To measure the objective difficulty and the reactions of the players to the character's adaptations (Hypotheses 1 and 3), we used the game logs. We used the *Game Experience Questionnaire* (GEQ) (Brockmyer et al., 2009; de Kort, IJsselstein & Poels, 2007) to measure players' perceived competence and challenge, and the perceived believability of the virtual character (Hypotheses 2 and 4).

The game logs kept track of the objective difficulty level during the game, which could be incremented at the moments for adaptation described in the previous section (see Table 12). Additionally, the game logs kept track of all the players' actions, allowing us to investigate which actions people chose after the virtual characters adapted his stance.

The GEQ provides several modules to determine players' experiences when playing and having played a game. All modules consist of lists which rely on self-report items players rate on a scale from 0 (*none/not at all*) to 4 (*extremely*). These modules are further divided into components that measure, among others, competence, challenge, and psychological involvement.

Three modules of the GEQ are of importance for this experiment: the main module (GEQ-core), the in-game module (GEQ-IG) and the social-presence module (GEQ-SP). Table 16 provides an overview of the structure of the GEQ, detailing

the components of the modules and example items, see IJsselsteijn, de Kort and Poels (n.d.) for a complete description of the GEQ. We chose the challenge and competence components from the GEQ-IQ module to give an indication of the competence of players and challenge of the game. These components consisted of two items each, resulting in four items which players were required to answer at several points during gameplay.

The GEQ-SP module consists of items that constitute three separate components: behavioural involvement, empathy, and negative feelings. Items constituting the empathy component indicate how the moods of participants and the virtual juvenile influenced each other. Items related to negative feelings indicate participants' dislike of the virtual juvenile. Behavioural involvement is measured through items that indicate causal relationships between the actions of participants and the virtual juvenile. We used items from this module to investigate players' relations to the virtual characters in the game.

We incorporated these measures in the game as follows. To determine the perceived difficulty of the scenarios, players were asked to answer the chosen four items from the GEQ-IG module to indicate how competent and challenged they felt at the end of each of the three scenarios. Having completed the game and having rated the final four items of the GEQ-IG, participants filled in the GEQ-core and GEQ-SP (see IJsselsteijn et al. (n.d.)). The final part of the experiment was a short demographics questionnaire, equivalent to the one used in the previous experiments with LOITER, see Section 6.4.

7.3.2 Results

In total, 42 participants participated in the experiment. One of them did not complete the final questionnaire and has therefore been omitted from the results. All participants were first-year students at the Dutch Police Academy, as we concluded from the results of our previous experiment (see Section 6.6) that LOITER may be too easy for more experienced police students. The mean age of participants was 26.1 ($SD = 7.5$); 56% was male, 24% female, and 20% did not report their gender. Participants indicated they had moderate experience with computer games ($M = 2.3$, $SD = 1.2$ on a scale from 0 to 4). Of the 41 participants, 14 completed the APJ condition, with both the adaptation and the justifications; 14 completed the AP condition, with just the adaptation; and 13 completed the control condition, without adaptations or justifications.

7.3.2.1 Hypothesis 1

For Hypothesis 1, we analysed how both the subjective and objective difficulty of the game changed over time. For the control condition, the objective difficulty always remained at 1, as the game never adapted its difficulty in that condition. Figure 40 shows how the average objective difficulty for three conditions varied over time, measured at the six moments of adaptation and the start of the game. At the first moment of adaptation (halfway scenario 1), the objective difficulty for all participants in both condition AP and condition APJ was incremented. Eventually,

Figure 40.: The change in objective difficulty over time for all three conditions.

participants in condition AP reached a difficulty level of, on average, 5.86 (out of 7, $SD = 1.351$). Participants in condition APJ reached a difficulty level of 6.14 ($SD = .663$). A repeated measures ANOVA (between-subjects factor: condition) revealed that the difficulty differed significantly over time, within subjects of both condition AP and APJ ($F(12, 228) = 95.934, p < .0005$). Post hoc tests using the Bonferroni correction did not reveal significant differences between conditions AP and APJ ($p > .05$).

7.3.2.2 Hypothesis 2

The scores of the GEQ-IG and GEQ-core give an indication of the subjective difficulty of LOITER-AP. Table 17 shows the participants' average scores of perceived challenge and competence for all three conditions and Figures 41 and 42 present a graphical overview of how these changed over time. We carried out ANOVAs with these scores as variables (fixed factor: condition) to determine whether there were significant differences between conditions for the challenge and competence after scenario 1, 2 and 3, and after participants finished the game. None of these ANOVAs revealed significant differences between the three conditions ($p > .05$).

Next, we investigated the perceived challenge and competence per scenario across conditions. A repeated measures ANOVA (within-subject variables: in-game challenge for scenario 1, 2 and 3; between-subject factor: condition) showed that there was a significant difference between the scores for perceived challenge ($F(1.982, 75.318) = 4.459, p = .015$). A post hoc test using the Bonferroni correction showed that this difference existed between the scores for scenario

Figure 41.: The change in perceived challenge over time for all three conditions.

1 and 3 ($p = 0.018$), but not between scores for scenarios 1 and 2, and 2 and 3 ($p > .05$). This means that, across all conditions, the perceived challenge of scenario 3 was higher than that of scenario 1.

A similar analysis with a repeated measures ANOVA (within-subject variables: in-game competence for scenario 1, 2 and 3) showed that there was a significant difference between the scores for perceived competence ($F(1.904, 72.358) = 5.170$, $p = .009$). A post hoc test using the Bonferroni correction showed that this difference existed between the scores for scenario 1 and 2 ($p = 0.015$), but not between scores for scenarios 1 and 3, and 2 and 3 ($p > .05$). This means that, across all conditions, the perceived competence of players while playing scenario 2 was lower than that for scenario 1.

These results indicate that, on average, players felt they had high competence, while they did not feel challenged. Figure 43 shows a scatter plot of the self-report ratings of players' competence and challenge. Comparing this to Vygotsky's ZPD (Figure 36), this scatter plot shows that participants may have been in a state of boredom while playing LOITER-AP. This implies that the game may still have been too easy to provide an apt challenge to players, even though the participants were first-year students.

7.3.2.3 Hypothesis 3

We investigated the logs of participants' play sessions which LOITER-AP stored to determine whether the juvenile invited participants to adopt a mirroring stance. Our analysis revealed that, on average, players most often chose friendly

Figure 42.: The change in perceived competence over time for all three conditions.

Table 17.: Participants' average scores for challenge and competence during and after the game.

MEASURE	TIME PERIOD	SCORE PER CONDITION (M, SD), OUT OF 4		
		C	AP	APJ
GEQ-core chal- lenge	Total game	.68 (.379)	.83 (.72)	.83 (.514)
GEQ-core com- petence	Total game	3.09 (.487)	2.69 (.655)	2.87 (.763)
GEQ-IG chal- lenge	Scenario 1	1.08 (.838)	1.39 (.656)	1.46 (.692)
	Scenario 2	1.27 (1.053)	1.79 (.848)	1.68 (.608)
	Scenario 3	1.31 (1.011)	1.79 (.802)	1.79 (.822)
GEQ-IG com- petence	Scenario 1	3.12 (.583)	2.82 (.541)	2.75 (.580)
	Scenario 2	2.88 (.795)	2.50 (.679)	2.57 (.703)
	Scenario 3	3.08 (.641)	2.79 (.378)	2.68 (.846)

Figure 43.: A scatterplot of the ratings of LOITER-AP's challenge and competence.

approaches to dealing with the juveniles. In total, approximately 91% of all player actions represented either a DF or SF stance. Table 18 shows the average frequencies of players' actions per condition. Four ANOVAs (dependent variables: number of DF, SF, SH or DH actions; fixed factor: condition) did not reveal any significant difference between conditions ($p > .05$). To determine whether players' behaviour was influenced by the adaptation in the game, we investigated their actions directly following an adaptation, in conditions AP and APJ. Of the 14 participants in condition AP, four responded at least once with a hostile stance (either DH or SH) to the juvenile changing his stance to a hostile one. In condition APJ, six participants reciprocated the juvenile's hostile behaviour at least once. The ratio of these responses compared to the desired, friendly responses when an adaptation was performed yields an average of 8.3% ($SD = 13.9\%$) hostile responses for condition AP and 15.7% ($SD = 21.3\%$) for condition APJ. An independent samples t-test on these variables showed no significant difference ($p > .05$).

7.3.2.4 Hypothesis 4

For Hypothesis 4, we investigated the participants' scores of the GEQ-SP, relating to their feelings of social presence. Table 19 shows, per condition, the average scores for each of the three components of the GEQ-SP: empathy, negative feelings, and behavioural involvement. Players' empathy with respect to the virtual juvenile was moderate. An ANOVA (dependent variable: GEQ-SP score on this component; fixed factor: condition) did not reveal any significant difference between conditions

Table 18.: The average frequencies of players' actions in LOITER-AP.

CONDITION	NUMBER OF ACTIONS PER STANCE, OUT OF 24 (M, SD)			
	DF	SF	SH	DH
C	10.15 (2.940)	11.62 (2.959)	1.08 (2.253)	1.15 (.987)
AP	8.14 (1.791)	13.36 (3.973)	.93 (1.859)	1.57 (1.555)
APJ	9.86 (4.240)	12.36 (4.909)	.57 (.646)	1.21 (1.251)
Total	9.37 (3.207)	12.46 (4.013)	.85 (1.682)	1.32 (1.274)

Table 19.: Participants' average scores for the GEQ-SP scales.

MEASURE	SCORE PER CONDITION, OUT OF 4 (M, SD)		
	C	AP	APJ
Psychological involvement, empathy	1.53 (.874)	1.74 (1.014)	1.74 (.578)
Psychological involvement, negative feelings	.82 (.580)	1.09 (.664)	1.10 (.469)
Behavioural involvement	1.47 (1.130)	2.12 (1.051)	2.44 (.778)

($p > .05$). Neither did an ANOVA for the second component, about negative feelings toward the virtual juvenile, reveal any significant difference between conditions (dependent variable: GEQ-SP score on this component; fixed factor: condition, $p > .05$). On average, players did not have negative feelings toward the loitering juvenile. The average of the players' behavioural involvement ranged from below to above moderate.

The scores for behavioural involvement differed significantly as determined by a one-way ANOVA (dependent variable: GEQ-SP score on this component; fixed factor: condition, $F(2, 38) = 3.280$, $p = .049$). A Tukey post hoc test revealed that there was higher behavioural involvement in condition APJ ($2.44 \pm .78$, $p = .041$) than in condition C (1.47 ± 1.13). There were no significant differences between conditions AP and APJ, nor between conditions C and AP ($p > .05$). The ratings for the behavioural involvement were on the lower half of the scale for the players in the control condition ($M = 1.47$, $SD = 1.130$), the addition of both adaptive difficulty and justifications in condition APJ yielded a significant increase ($M = 2.44$, $SD = .778$), raising this rating about one point on the four-point scale.

7.3.3 Discussion

The results entail the following answers to the research questions. Adaptivity made LOITER-AP's objective difficulty rise in both adaptive conditions, with no significant differences. On average, players reached the second-highest difficulty level. This indicates that they performed well in getting the juvenile to cooperate,

otherwise the difficulty would not have been incremented. Looking back to the experiment with LOITER-FB in the previous chapter, this further supports the conclusion of that game having a ceiling effect where the players simply could not do better, as all players experienced increases in difficulty in LOITER-AP. We may conclude that, despite adaptive difficulty, this ceiling effect is still in place in LOITER-AP, as players performed almost as well as possible. The findings about the game's subjective difficulty agree with this: players felt the game was not challenging, while at the same time they felt they were competent. Comparing the scatterplot of the players' ratings of challenge and competence (Figure 43) to Vygotsky's Zone of Proximal Development (Figure 36), we can see that players may indeed have felt bored while playing the game. We can conclude that neither the addition of adaptive difficulty nor justifications had an impact on the players' ratings of challenge and competence, as compared to the game without these features. While LOITER-AP incremented its difficulty based on players' competence, the implemented improvements did not increase the game's challenge, despite the participants having less experience with social interactions in the police domain than the participants of the previous experiments (see Chapter 6).

Players of LOITER-AP felt neutral towards the virtual juvenile; they did not indicate they felt empathic with him. This also shows in their ratings of negative feelings towards the juvenile, as these were also on the low-end of the scale. The game logs reflect the same finding: players mostly took a friendly approach, not taking up the invitation to become hostile. This held for all three conditions; LOITER-AP's adaptation did not change players' negative feelings or feelings of empathy towards the virtual juvenile. However, there is an important finding concerning the behavioural involvement of players. The addition of both adaptive difficulty and justifications yielded a significant increase of approximately one point on the four-point Likert scale, as compared to the control condition which lacked these additions. In condition AP, which included adaptive difficulty, but lacked justifications, this increase was also visible, but not significant. This may mean that the justifications provided the extra information necessary for players to interpret the juvenile's behaviour as more involving. This might have resulted from the more complex or more coherent nature of the juvenile's behaviour. This finding entails several possible conclusions. It may be that the juvenile's switching of stances provides a richer interaction: he no longer simply follows the rules of the interpersonal circumplex, but chooses a different course of action. This may have drawn players' attention as it provides an unexpected experience, thus letting them feel more involved with the juvenile as they had to keep paying attention to what he would do next. Nonetheless, the increase of behavioural involvement is only present when the juvenile's justifications for his behaviour are shown. This leads to the conclusion that this increases the perceived causality of the juvenile's behaviour and therefore of the game's coherence.

Overall, players did not opt to take hostile actions very often in LOITER-AP. From this, we conclude that players do not necessarily feel invited to reciprocate the juvenile's adoption of a hostile stance. While the results of the GEQ-SP indicate that players felt stronger behavioural involvement with the juvenile in condition APJ, they did not let themselves be provoked to act hostile. This may be a result of

participants' feeling of presence being too low, which is reflected in the scores for the GEQ-SP components about empathy and negative feelings. However, why the participants mainly chose friendly actions is unclear. It may be that participants understood the mechanics of Leary's Rose and intentionally continued taking a friendly stance, no matter how unfriendly the juvenile acted. Alternatively, participants may have recognised that taking a friendly action would always end in a desired ending, without intentionally reflecting on Leary's Rose.

As discussed in the beginning of this section, LOITER-AP was still prone to several shortcomings in its design. The updated design of the game entailed a more complex cognitive model for the virtual juvenile which had a positive effect on participants' involvement. However, the mechanics of the game still allowed most users to solve the three scenarios satisfactorily without them feeling challenged. As in LOITER-TB and LOITER-FB, players' choices are always limited to a given set. The way these choices are presented in LOITER-AP could have made the game pretty straightforward to complete without the necessary challenge. For future iterations of LOITER, multiple shortcomings need to be addressed. First and foremost, the game needs to become more challenging. LOITER's difficulty can be increased in several ways. Rather than simply having one type of player behaviour which is always correct, different situations in the game can call for different approaches, for example, being more friendly one time than other times. To challenge players to put more thought into their decision process, they may benefit from a different input than a multiple-choice format. When players voice their own ideas about what to do, they are primed less and have to think more intently about which course of action may be the best. As the results of this experiment show, the expansion of the juvenile's cognitive model improved a key factor of the game: behavioural involvement. The juvenile seemed more real when he changed his behaviour and did not simply reciprocate the players' behaviour. However, players of LOITER-AP did not yet feel invited to mirror the juvenile's behaviour, for example, by becoming hostile towards the juvenile when he mocked them. This may be due to a lack of proper realism, which may be alleviated in various ways, for instance with richer behaviour with more variability. Improved fidelity of the game, such as in a virtual reality (VR) setup, may also improve players' presence. In Chapter 9, I discuss the first steps taken to implement the principles of LOITER in such a VR environment.

7.4 REFLECTIONS ON THIS CHAPTER

In this chapter, I addressed the fourth Research Question of this dissertation:

RQ4 How can virtual agents use meta-thinking to adapt serious games to the user's competence?

To answer this question, I discussed how serious games can benefit from the adaptation of their challenge to a player's competence. The theory of the Zone of Proximal Development (Vygotsky, 1978a) is highly influential to this approach, as is a key factor of the flow theory of Csíkszentmihályi (1990), namely the matching a task's challenge to a person's competence. Several systems have been designed

to implement this concept, ranging from games that adapt to players on a macro-scale, for example, by ordering their missions depending on players' behaviour, to games that adapt on a micro-scale, adapting small parts of the game world such as the provision of items, or by providing hints.

For the design of LOITER-AP, a serious game prototype for social skills training, I chose a micro-scale approach which uses adaptive profiles. This approach is inspired by the characters of the Virtual Storyteller reasoning out-of-character (see Chapter 4). The NPC of LOITER-AP can change his stance profile depending on how well players perform in the game. For example, when a player is able to convince the juvenile to cooperate, the game interprets this as that player having increased their competence. The juvenile is not aware of this knowledge in-character as this knowledge is only available out-of-character. Reasoning on this meta-level, the game adapts the juvenile's profile so that he is inclined to be less cooperative with the player, incrementing the game's challenge. Because this change in the juvenile's behaviour may be interpreted by players as incoherent, LOITER-AP provides justifications to retain coherence by showing the events causing this change, which either involved misinterpreting players' actions or 'making up' excuses. These justifications are shown through the meta-techniques of thought bubbles and flashbacks, as discussed in the previous chapter.

An evaluation with students of the Dutch Police Academy showed that most players of LOITER-AP were able to complete the game successfully, ending with the highest levels of difficulty. This result, combined with the game being rated as not very challenging, indicates that they found the game, on average, not difficult. Several limitations of the game may have caused this result. For example, players chose actions in a multiple-choice fashion, which may have made it too easy for them to determine which was the correct response to the juvenile. Moreover, the cognitive model governing the juvenile's behaviour relied on the theory of Leary's Rose (Leary, 1957), with the exception of the adaptations. The incorporation of other models, for example, those relating to *rappport* and *face* (see Section 5.2.1), may positively influence the game's perceived realism while also increasing its difficulty. Changes in these game mechanics may improve the challenge of the game, requiring more effort on the player's behalf to think on which approach to take. A next step in making the experience of such a game for training social skills more immersive also lies in the use of higher fidelities for the training environment, for example, through the use of virtual reality (see Chapter 9). This allows players to experience interactions in a more immersive way, taxing them more to take correct decisions under stress.

Nonetheless, the results of the evaluation also showed that the addition of justifications, combined with the adaptations of the juvenile's profile, resulted in improved behavioural involvement of players. This highlights how the meta matters when adapting the challenge of a serious game: the use of adaptive profiles and justifications leads to the apparent coherency of the behaviour of virtual agents.

PART IV

CONCLUSION: REFLECTING ON THE RESEARCH

In this part, I reflect on my research.

In Chapter 8, I answer the four research questions I posed at the start of this dissertation and discuss how the meta matters for interactive storytelling and serious games.

In Chapter 9, I close with an outlook on future work. I discuss the possible improvements to the Virtual Storyteller and the Interactive Storyteller, and describe our first foray in implementing LOITER in a virtual reality environment.

Vlad was right. There are no choices. Nothing but a straight line. The illusion comes afterwards, when you ask 'Why me?' and 'What if?' When you look back, see the branches, like a pruned bonsai tree, or a forked lightning. If you had done something differently, it wouldn't be you. It would be someone else looking back, asking a different set of questions.

Max Payne

Max Payne

Remedy Entertainment (2001)

THIS chapter provides conclusions from the research discussed in this dissertation. In Section 8.1, I revisit Chapter 1, recapitulating how previous work on meta-thinking has influenced this dissertation. In the following two sections, Section 8.2 and 8.3, I answer the four research questions related to meta matters in interactive storytelling and serious games, respectively. Based on these answers, I go meta myself in Section 8.4 and discuss what this paradigm offers for research on interactive digital technology.

8.1 THE META-NARRATIVE OF THIS DISSERTATION

To reiterate what I said in the introduction of this dissertation: stories help us understand the world we inhabit. The goal of this dissertation's narrative is to understand how thinking about stories and games can augment our experiences with interactive storytelling systems and serious games. This dissertation chronicles my efforts to answer the following question:

How can both users and virtual agents use meta-thinking to enrich interactions with interactive storytelling systems and serious games?

As I explained in the introduction, experiences with such systems can be enriched in different ways to improve the enjoyment of stories or increase the efficiency of a serious game. For example, work can be done to improve the visual fidelity of virtual environments to make them more life-like, or to expand the player's input modalities with virtual reality. My approach to improving interactions with interactive storytelling systems and serious games is aimed at understanding the relation between the world of play and the external world. The starting point for this is the well-known concept of the magic circle, which functions as a distinction between real and virtual worlds (Salen & Zimmerman, 2004). I oppose the view of Caillois (2001), who asserts that there is a strict boundary between the two that cannot be crossed while playing. This dissertation

revolves around the idea that the ability to actually cross this boundary is of paramount importance to enriching experiences of interactive storytelling systems and serious games. Huizinga's classical work *Homo Ludens* (Huizinga, 1949) as well as *Man, Play and Games* by Caillois (2001), and *The Hero with a Thousand Faces* by Campbell (1949) state that both readers and players knowingly enter a fictional world when they read and play. While Caillois did not agree with players being able to be present in the virtual and real world simultaneously, Bates et al. (1992) point out why it is important that players can transcend their existence in the virtual, fictional world. For example, when children are playing, but one of them gets hurt, the other is able to say 'You are dead.' As games take place in the physicality of the real world, players are able to step out of the game world to reflect on their experience and, for example, stop playing, or decide to change their play in a way that suits them.

This also holds for experiencing stories: readers reflect on their reading process, based on personal preferences and expectations of stories. They do not read a story in isolation from the real world, but let the world external to that of the story influence their reading. For example, one might be more inclined to read thrilling, action-packed page-turners, while the other enjoys a slowly developing, metaphysical drama. It may also be the case that the reader wants to finish a chapter quickly, because she has to get off the train at the next stop. Especially those stories that are packed in metaphors and analogies ask readers to read between the lines and think about the possible allegories and allusions. Thus, going meta while reading stories enables readers to enrich their experiences.

In this dissertation, I describe how going meta can enrich the experiences of players when interacting with interactive digital storytelling systems and digital serious games. My research involved investigating interactions with both the *Virtual Storyteller* (VST), an interactive storytelling system, and a series of serious game prototypes called LOITER. In both these types of systems, players interacted with virtual agents: characters in a fantasy domain in the VST, and virtual juveniles in law enforcement scenarios in LOITER. I approached the challenge of going meta from two sides: from that of the players of these systems and from that of the virtual agents with whom players interact in these systems. Both these entities inhabit the virtual worlds of either system, yet both can also take a meta-perspective and reflect on the experiences of the other party and of themselves. Through reflecting and reasoning about the state of the virtual world and the state of the other entity, they can adapt their behaviour to improve their interaction through and with the system. In other words, any type of agent, whether human or artificial, can benefit from adopting this perspective. In Parts II and III of this thesis, I discussed my research on the Virtual Storyteller and LOITER, respectively. In the following two subsections, I recapitulate my findings in the same order.

8.2 INTERACTIVE STORYTELLING GONE META

To investigate to what extent the meta matters in interactive storytelling systems, I analysed the VST from the point of view of players and of virtual agents. My research continued the work of Swartjes (2010) about the inner workings of the

Virtual Storyteller. For my research, I focused on the user experience of players (Section 8.2.1) and the plotting capabilities of agents (Section 8.2.2).

8.2.1 *How players go meta to co-create stories*

In Chapter 3, I sought to answer Research Question 1 of this dissertation.

RQ1 *What kind of meta-thinking do people do while co-creating stories?*

To answer this question, I investigated interactions with the *Interactive Storyteller* (IST), which is based on the VST. The VST was originally meant to generate stories, with the Interactive Storyteller adding a layer on top of it to make it interactive (Alofs et al., 2015). Interacting with the IST happens through a graphical user interface with multi-touch capabilities so that players can select actions for one or more characters and move them around by dragging them. Other characters act autonomously when not controlled by players. A study done by Alofs et al. (2015) showed that pairs of children who interacted with the IST were actively discussing among themselves what to do. In other words, they went meta to co-create stories. In Chapter 2 of this thesis, I describe how we continued this line of research by analysing the interactions recording during Alofs et al.'s study. We gained an understanding of how children go meta when interacting with the IST by analysing their interactions in two iterations.

Firstly, we analysed the story logs to investigate the causality of the created stories. We found that the children who participated in this study were able to causally connect events happening in the story with logical reactions, making the stories they co-created locally coherent. We also saw that the children created stories with clear goal-action-outcome sequences, indicating they were able to create globally coherent stories. However, this coherence was only partial, as children also experimented with the possibilities of available actions and their effects. This partial coherence is due to almost half of these actions not affording children to carry out new actions, instead simply providing an effect without consequence in the story. The children seemed to have mixed intentions: on the one hand, they did their best at advancing the story, on the other, they played around with the IST to explore its affordances. Only looking at the logs of stories the children produced did not give us sufficient information to determine why they took certain actions. We hypothesized that factors not visible in the story logs were the cause of certain events happening. Therefore, we investigated the children's meta-communication about their interactions in a follow-up study.

In our second study, we further analysed the children's communication by looking at the video recordings. This turned out to be essential to understanding several cause-action sequences in the created stories. The children actively communicated with each other to deliberate which actions to take and how to plan the desired outcome of a story. More than simply copying character actions, they also went meta on several levels. For example, the children indicated to what extent they liked the story and asked each other questions about the mechanics of the interaction. The IST has similarities to old-fashioned board games, both in appearance and in use, for example, because of the layout of the story world

and because of its turn-taking mechanic. This may explain why children at times treated the IST like a board game, indicating whose turn it was and what they would do in later turns. Additionally, the children sometimes complained about their interactions, being confused by the user interface.

We expanded on work by Sawyer (2003) and Fine (2002), who distinguish between different frames for interaction, by creating the Perspective $\times$ Reference ($P \times R$) annotation scheme. This 3×3 scheme can be used to classify the communicative acts of players with interactive storytelling systems such as the Interactive Storyteller. Any communicative action of a player can be uttered from either the story, play or real world frame (*perspective*). Such an action can refer to either the story world, the game system, or the real world (frame of *reference*), resulting in nine different classes. These classes helped us to understand how the children approached the IST at different moments during their interactions. We discovered that children indeed actively go meta, not at all confining themselves to the story or play frames. For example, they talked about the realism of the behaviour of the virtual characters and they explained their behaviour to their partners so that they could collaborate.

Our conclusion is that, by analysing the combination of the children interactions with the Interactive Storyteller and also with each other, we can sketch a clearer picture of what happens when people interactive with storytelling systems. To children, going meta matters when co-creating stories. This research has been a first step at discovering how the meta matters to them with the end goal of improving interactive storytelling systems. With the help of the $P \times R$ annotation scheme, future research can provide insight into how players of other interactive storytelling systems interact with these systems.

8.2.2 How agents go meta to perceive, assume and plot stories

Turning to the side of virtual agents in the Virtual Storyteller, Chapter 4 deals with our work on agents that can go meta to improve stories. This chapter addressed Research Question 2.

RQ2 How can virtual agents use meta-thinking to improve the generation of stories?

All stories create an experience for their readers. In the vST, this is done with virtual characters which adopt goals to pursue. Their antics in the story world with each other result in stories. The author has a vital role in this process, as she decides what the story world looks like, which actions characters can perform and which goals they can adopt (Swartjes & Theune, 2009a). On the one hand, authoring the story worlds in such a way that only specific stories can occur limits the possibilities of the virtual characters and the creativity of the interactive storytelling system as a whole. On the other hand, giving the virtual characters too much freedom in choosing what they can do may lead to dull stories. To maintain a balance between these two extremes, virtual agents can go meta about the story to determine what they should do. Similar to the agents in work by Aylett et al. (2006), virtual agents in the vST have both an in-character (ic) view of the virtual world and an out-of-character (ooc) view. The former is their character role in

the story: on this level, they should only perceive what the character senses to maintain realism. The latter view is their meta-level which lets them have an omniscient view of the virtual world, enabling them to reason about the story. In previous versions of the VST, the character agents were omniscient, which made particular types of stories impossible. For example, stories about deception depend on characters not being omniscient. We applied the IC/OOC paradigm to enable agents to create such stories in the VST.

Firstly, we implemented actions for the character agents that restrict them to sense only a part of the world they are currently inhabiting. No longer being omniscient in their in-character perspective, the virtual agents cannot access all the information about the story world, for example, where an item is located they are looking for. As a side-effect, this prevented characters from making plans to achieve certain goals, as they could not reason towards a solution without information they could not perceive. Secondly, we modified the planner of the VST to allow for agents to use assumption operators to assume facts, for example, assuming an item is located at a specific location, even if it is not. When the character has to plan towards a goal, it immediately goes meta, into its out-of-character perspective, to get a complete overview of the story. Based on this omniscient world view, the virtual agent can then self-deceive his character role to stimulate the creation of more dramatic stores. For example, the item the character is supposed to be looking for is located at a certain location. However, instead of simply planning to move to that location and retrieve the item, it assumes, in-character, that the item is located at a different location. Its reason for this is that it knows, OOC, its enemy is located there. Thus, when it plans to go there IC, it knows OOC that it will be confronted with its enemy, leading to a conflict between the two of them. This conflict might otherwise have been avoided by the character, had it simply moved towards the item and retrieved it.

This is a proof-of-concept of what is possible when virtual agents go meta, reasoning in ways similar to those of improvisational theatre players. Going meta allows for these agents to do things they would normally not consider from an in-character point of view. For example, story characters may perform mental leaps of faith to arrive at conclusions that advance the plot, or they may choose to 'forget' about some property of an object to dramatic effect, such as lighting an exploding cigar. With proper authoring for the assumption operators, offering virtual agents to go meta can enrich the variety of generated stories and, in turn, the users' experience.

8.3 SERIOUS GAMING GONE META

In the second part of this dissertation, I explored how to use the concept of meta-thinking to design serious games for training social skills. I created several prototype games called LOITER, which features virtual characters with which players need to interact.¹ All scenarios in these games took place in the domain of law enforcement, with players taking on the role of a patrolling police officer. In these scenarios, players come upon a group of loitering juveniles who they have to

¹ In collaboration with T-Xchange, see <http://www.txchange.nl>.

convince to cease their loitering. Players can take different approaches to solving these conflicts. For example, they can choose to be more dominant and less friendly, or more affectionate and accommodating. The leader of the juveniles will respond to players' actions following the theory of interpersonal reflexes (Leary, 1957). All scenarios end with the juveniles ceasing their loitering, yet whether they do so willingly or under pressure affects how they interact with the player in successive scenarios, thus influencing long-term interaction.

In this dissertation, I describe how the meta matters when designing such serious games. Similar to my investigation in interactive storytelling systems, I approached this from the perspective of players and from the perspective of agents. For players, this means going meta to reflect on their learning experience (Section 8.3.1); for virtual agents, using information on players' competence to adapt their behaviour (Section 8.3.2).

8.3.1 *How to support players to meta-think and reflect*

In Chapter 6, I sought to answer Research Question 3 of this dissertation:

RQ3 *How can serious games stimulate meta-thinking to increase their effectiveness?*

What distinguishes serious games from regular games is that the former have the explicit goal to teach players something which is of importance outside the play world. This entails that knowledge gained by playing a serious game needs to be understood by players outside of the context of the game itself. In other words, players need to translate this knowledge to real-life scenarios, reflecting on their gaming experience and becoming aware of how it relates to the real world. As such, effective serious games require players to go meta to learn.

For LOITER, I designed two meta-techniques to stimulate this process of reflection and thus increase the game's effectiveness (see Chapter 6). The first of these is the use of thought bubbles which give insight into the juvenile's thoughts about a player, showing the latter's stance. The second is the use of flashbacks which indicate how the juvenile remembers the player's behaviour in a previous scenario. Each flashback shows an action the player took which corresponds to his average stance during that scenario, together with an interpretation of the juvenile and an explanation of how the player's past behaviour will cause the juvenile to act in the upcoming scenario. Both of these techniques were intended to make the cause-effect relation between players' behaviour and the juvenile's responses more explicit.

We evaluated these games with several classes of students from the Dutch Police Academy, investigating to what extent the addition of these meta-techniques influenced their learning experience. The results indicated that learning gains of the games with meta-techniques were, on average, minimal to non-existent. Neither were there learning gains in the games which lacked the meta-techniques. There are several possible explanations for this. First and foremost, the meta-techniques may not have achieved their intended goal of making players reflect on their behaviour. Players may not have been aware of the use of these techniques, possibly not understanding the information contained in them or

simply ignoring them. This calls for closer scrutiny of the type of information that they may benefit from in such a serious game. Secondly, the game's overall design influenced multiple factors relating to the players' experience. Each scenario was straightforward; there was limited branching of the story. All players took the same number of actions, with each scenario ending with the juveniles leaving. Moreover, the actions available to players were presented to them in a multiple-choice fashion: verbal utterances were accompanied by pictures of the player character's posture that matched those utterances. This may have resulted in players almost always choosing actions that resulted in positive reactions from the juvenile, making the game too easy to complete successfully. Thirdly, there may have been a ceiling effect in the students' learning process, which may have caused the lack of a difference between learning gains in students who played the games with versus those who played without meta-techniques. This may in part be due to the students already having had experience with situations similar to those in the game, as they were in the third year of their curriculum. Fourthly, the games may also have been too simple to convey the complex social interactions of real life, thus being too distanced from the real world for students to translate their learning gains from the game to the real-world settings in the situational judgement tests. It may also have been the case that these tests did not correctly replicate the actual real-world settings.

While the addition of meta-techniques did not improve learning gains, these results emphasize that designing serious games is a process which entangles many requirements. Learning objectives are the core of any serious game, yet serious games should also be entertaining, otherwise, the main motivational aspect of such a game is lost. Furthermore, serious games need to provide a challenge which suits the players that play it. In the studied LOITER games, this challenge was too low, as explained above. To overcome this hurdle, I investigated how the virtual agents in LOITER can adapt the game's difficulty to match a player's competence.

8.3.2 *How agents go meta to adapt a game's difficulty*

Chapter 7 is dedicated to Research Question 4 of this dissertation:

RQ4 *How can virtual agents use meta-thinking to adapt serious games to the user's competence?*

Based on the results of our previous studies on the LOITER games, I decided that several elements of the game design should be changed to provide a better challenge to players. The core mechanic of these games is the choice players have in which approach to take when addressing a loitering juvenile. His responses, and hence the outcome of a scenario, depend on these choices. Chapter 7 of this thesis reaches back to our research on agents going meta in interactive stories to stimulate conflict between characters. This is the crux of adaptive difficulty in a revised version of LOITER. I decided that, to match the challenge of the game to the competence of a player, virtual agents should adapt their behaviour depending on that player's competence in the game.

Like the virtual agents in the VST who went meta to seek out conflict with other characters, the virtual juvenile in LOITER can go meta to choose to be less or more inclined to cooperate with players. This is what I termed a character's *adaptive profile*. To adapt his profile, the virtual agent reasons out-of-character by analysing whether a player has been taking the correct actions to kindly convince the juveniles to cease their loitering. When a player chooses the right options to de-escalate the conflict, the juvenile can choose to become less cooperative, resulting in him being more difficult to convince to go away in a friendly manner. This may seem counter-intuitive to players, as they had been taking a friendly approach in dealing with the juvenile, yet, suddenly, this has an opposite effect: he becomes less friendly. To keep the juvenile's behaviour coherent, the agent has the ability to explain its in-character behaviour through *justifications*. These are shown through the meta-techniques of thought bubbles and flashbacks as discussed in the previous section. At set moments during scenarios, the juvenile evaluates the player's competence and decides whether or not to adapt his profile. When he decides to adapt during a scenario, he offers an explanation through a thought bubble by showing how he misinterprets a player's actions, for example, complaining about how the player is too much of a nice guy. When the juvenile decides to adapt his profile at the start of a new scenario, he explains his change in behaviour by showing a flashback which relates to some event that occurred 'off-screen', outside the context of the in-game scenarios. For example, he thinks back to his phone getting confiscated by his parents, which makes him less likely to collaborate with another figure of authority.

Alternatively, if the player seems unable to take the right actions, the juvenile does not change his adaptive profile. Lowering the difficulty of the game would entail the juvenile adapting a profile that would cause him to cooperate more easily. I chose not to implement this, because there was an apparent ceiling effect in the players' performance in the game, with the game being too easy for most players.

Similar to the previous studies on LOITER, we evaluated this design with several classes of police students of the Dutch Police Academy. Different from the participants in our previous studies, these students had only been enrolled for one month in the training of the Academy. The evaluation showed that, again, the learning gains of the students were minimal, and that the addition of the adaptive profile of the juvenile did not result in significant differences in learning gains. Additionally, neither the profiles nor the addition of justifications affected the players' approaches to dealing with the juveniles. However, there was an indication that players felt increased behavioural involvement with the juveniles when they played LOITER with adaptive profiles and justifications. This may mean that insights into explanations of the juvenile's behaviour can positively affect the extent to which players can relate to him. This was not apparent in the previous studies which also featured such meta-techniques. It may be the case that such meta-techniques are of added value only in more complex situations. For example, in the game used in this study, the interactions between players and virtual agents are not as clear-cut and straightforward, as juveniles suddenly showed major changes in their behaviour instead as opposed to behaving fully predictably. Therefore, the meta can matter when social interactions with virtual

agents become increasingly complex. This can happen when the agents themselves go meta and reason to behave differently, or when games afford richer possibilities for interaction with such virtual agents.

8.4 THE META-PARADIGM

In this dissertation, I investigated the design of and interaction with the Virtual Storyteller and LOITER to answer the main research question first asked in Section 1.3:

How can both users and virtual agents use meta-thinking to enrich interactions with interactive storytelling systems and serious games?

I took the perspectives of both players and virtual agents who interact with each other in these systems, looking at how they could use meta-information to improve user experiences. The reason for the use of meta-thinking is that virtual worlds in which games and stories take place are not strictly separated from the real world: they are an inherent part of it. The magic circle in which play takes part is not inescapable; on the contrary, when people play, they actively go meta, using their thoughts about their playing to inform decisions within the magic circle. This distinction between in-character and out-of-character thinking applies to both people and virtual agents: both can use out-of-character information to adapt their in-character choices. Revisiting Figure 44, it is clear that players and agents are, in the end, similar in this perspective. The *agens reputans* is an entity which is aware of the existence of the real world outside of a virtual world and can reason about it. This dissertation has provided various investigations into how different types of entities go meta or can be stimulated to go meta. My research serves as a first exploration of how the meta-paradigm can inform the design of interactive storytelling systems and serious games. It has made way for new opportunities for further research, such as how the intelligence of virtual agents can be augmented to take into account meta-information, and investigating why people take on different perspectives and how they may be stimulated to do so. In Chapter 9, I address two possible research directions related to the domains of storytelling systems and serious games.

In this dissertation, I attempted to model the interactions between human players and virtual agents. Taking on a meta-perspective, I link the virtual world and the real one using the concept of the magic circle. The virtual world is indeed possible by both humans and virtual agents. Going meta about these interactions myself allowed me to understand how both types of agents can go meta. Reflecting on how and why processes such as these occur is paramount to understanding them. And, in the end, this is exactly what research is about: going meta to model the world.

Figure 44.: Player-agent interaction. Both entities follow the same sense-think-act cycle and have the means to think on a meta-level about their interaction.

As the unknowable future becomes the unchangeable past, risk must collapse into certainty, one way or another.

John Nately
Into Darkness
Egan (1995)

THE research presented in this dissertation addresses how both virtual agents and human agents can take a meta-perspective on interactions with interactive storytelling systems and serious games. Besides addressing how this paradigm can assist these entities in improving such interactions, this work also opens new lines of enquiry and raises new questions as to how the findings presented here may affect other domains. In this chapter, I discuss possible continuations of research on the Virtual Storyteller and other interactive storytelling systems (Section 9.1), and the implications of creating a virtual reality serious game for social skills (Section 9.2). I close this chapter by sketching an image of the future for interactive storytelling and serious games in Section 9.3.

9.1 IMPROVING INTERACTIONS WITH THE VIRTUAL STORYTELLER

As Swartjes (2010, p. 190) notes in the closing chapter of his dissertation, ‘a system like the Virtual Storyteller is never finished.’ As much as Swartjes’ work was a continuation of earlier research by Faas (2002), and Theune, Rensen, op den Akker, Heylen and Nijholt (2004), so is the research in this dissertation a continuation of his. Similarly, the research presented in this thesis lays the basis for more investigations into how the meta matters to interactive storytelling.

In the previous chapter, I drew conclusions from my research, stating that going meta can indeed help researchers gain insight into interactions with ISSs. However, my research knew several limitations, due to the design of both the VST and the performed evaluations. Below, I address these and explain how future research could attempt to tackle them.

One prominent limitation of the VST is concerned with the agency of players. As is inherent to the design of interactive systems, virtual worlds always have their limitations. Virtual environments have their boundaries in time and space: they can only contain a limited amount of objects and characters to interact with, and only a limited variability in the range of actions players can perform. The latter is evident in the design of the VST: available actions are presented to players in a multiple-choice fashion. This is a strict way of enforcing players to choose what designers of these systems decided should be possible for them to do. Rather than limiting the players in this way, a different design could offer more freedom of

choice. For example, letting players speak sentences they would like their character to utter would allow players to say anything they want. This approach was applied in *Façade* (Mateas & Stern, 2003, see Section 2.1.3.2). To enable virtual characters to respond logically, the system then needs to interpret what a player says and means, which is a daunting task (Paris, Swartout & Mann, 2013). An even more daunting task is ensuring coherency between player actions and intended effects, not all the actions players take may be modelled to have the effects desired by those players. Moreover, players may take actions that deviate from the storyline intended by the authors of the virtual environment (the *narrative paradox*). The approaches to handling the variety of players' behaviours discussed in Chapter 2 mostly revolve around ensuring this coherency by linking an event available in the virtual world to a player's action through some form of justification. Like the justifications in LOITER (Chapter 7), this approach could also work in the vST, as an expansion of the current use of *late commitment* operators.

Another limitation of the current design of the vST is that virtual agents can only reason about the intentions of other virtual agents. Here, the challenge is to let virtual agents go meta and reason about the intentions of the players, for example, to determine which goals the latter are trying to accomplish. In the architecture of the vST, this can be done by investigating whether players' actions were intended to be causally related to other actions and events. One approach to solve this problem is found in Ha, Rowe, Mott and Lester (2012), in which player goals are recognised to adapt a serious game to their playing style. Our analysis of children's interactions in Chapter 3 showed that we can determine this when additionally taking the out-of-character behaviour of players into account. By automating this recognition of player behaviour, virtual agents may gain access to information leading them to understanding the players' intentions. The communication between players may, for example, be transcribed by a module for automated speech recognition (Benzeghiba et al., 2007). This transcription can then be processed to distil intentions using an ontology such as DIT++ (Dynamic Interpretation Theory, see Bunt (2009)) to classify utterances that may indicate the perspective taken by players when discussing certain frames of the interaction with the Interactive Storyteller. Uniting the virtual and physical in this way may lead to a step forward in understanding the players' intentions. With this understanding, the virtual agents may take the players' plans into account to create more compelling stories, for example, by doing the inverse of what players expected.

While they are playing with the vST, children are limited in their co-creation process by the content the system offers. Interaction with the vST could be improved by expanding these co-creational affordances. Instead of providing a pre-specified virtual world to players, they could be authors of the story world themselves. This implies that the children should go meta in advance of playing their story, as they need to reason about the function of characters, items and locations in the storyworld. Their design choices would limit the story world to something that is bounded, but comprises the story elements that fit that story. Additionally, through combining different types of these elements, children may already come up with original story worlds.

These are but a few of the possible improvements from which the Virtual Storyteller could profit. The ideas presented in this section could be extended to other interactive storytelling systems: all such systems would benefit from gaining more insight into the types of players and their intentions. Player agency is the crux of interactive storytelling; it is the player who decides what she wants to in an interactive story. Therefore, more research should focus on understanding players' desires when interacting with interactive storytelling systems. Plenty of research has focused on increasing the artificial intelligence of virtual agents to adapt to player actions (Yannakakis & Togelius, 2014). Still, in most commercially available interactive stories and games, player choice is limited to choosing one of a limited number of available actions to take. By delving deeper into the player's mind, we may be able to understand in more detail what she wants from such a system. With this information, we can adapt the design of interactive storytelling systems to better cater to the player's wishes, improving her experience.

9.2 LOITER-VR: SOCIAL SKILL TRAINING IN VIRTUAL REALITY

The LOITER series of serious games was intended to improve players' social awareness. The discussed implementations focus on making players reflect on their approach, assisted by meta-techniques. The visual design of these implementations contains static imagery: the virtual characters do not move; instead, their appearance only changes when their stance towards someone changes. In the context of the overall design of the game, these animations fit the characters' behaviour, which is relatively simple.

Figure 45.: A player and his first-person, in-game view of LOITER-VR.

To investigate how the same educational objectives would work in a *virtual reality* (VR) environment, we worked on a prototype called LOITER-VR (Linssen & de Groot, 2013). Figure 45 shows a side-by-side view of a player wearing the immersive VR suit which registers his movements, and his view on the virtual world. Like the other LOITER games, this version features a scenario in which the player tries to convince a loitering juvenile to cease his loitering. This serious game builds on the concepts used in the VST and LOITER, with the juvenile responding to the approach a player takes. For the virtual environment, we collaborated with RE-

VIRTUAL
REALITY (VR)
LOITER-VR

liON, a company specialized in offering tactical training simulations through multi-modal VR.¹ We tackled four design issues to create a working prototype of LOITER-VR. Firstly, we designed an architecture to link our software, which contains the cognitive models of the virtual characters, to the virtual environment, which contains the visualized characters. This was done through a message broker system which enables software modules to write on and read different topics, allowing them to retrieve information from different sources. Secondly, we created an ontology which related abstract facts about the story world, for example, locations and objects, to their visualised counterparts. This allows the virtual characters to reason about the objects in the virtual environment. Thirdly, the modular setup of our architecture allows other modules to listen in on the information being exchanged. When a trainee is using LOITER-VR, a separate visualisation allows a trainer to follow the actions of the trainee and the virtual characters. Moreover, this visualisation allows the trainer to influence the behaviour of the virtual characters, acting as a director.

The fourth design issue was the intelligence of the virtual juveniles. In LOITER-VR, the juveniles use the same behavioural model as in LOITER-TB, which is based on Leary's theory of interpersonal reflexes (Leary, 1957). Based on the player's actions, the juvenile can respond differently. A major difference between the 2D LOITER games and LOITER-VR is how player input is handled. In the former, four options are presented to players. In the latter, there is no overview of available actions: players have to decide for themselves which action to take. This results in there being more player agency than in LOITER-TB, which makes the game more challenging than the previous LOITER games. However, this also introduces the hurdle of recognizing and interpreting player behaviour to give input to the virtual juveniles so they can perform a suitable response to the player. Due to our focus on the interaction itself, we did not implement an automated approach to take this hurdle. Instead, we used a Wizard-of-Oz setup in which a domain expert interpreted the player's behaviour in terms of the behaviour model used by the virtual juveniles. This interpretation was then processed by the cognitive model of the virtual juvenile to decide which of the responses available to him was the most fitting. Automating the recognition of player behaviour poses several possibilities for future research. Both the verbal and nonverbal behaviour of the player should be sensed before being interpreted. As described in the previous section, interpreting players' intentions may be done using a taxonomy such as DIT++ (Bunt, 2009). In the context of LOITER-VR, player behaviour should be mapped onto the interpersonal circumplex as proposed by Leary (1957), yet which is difficult to do automatically as shown by Vaassen and Wauters (2012). For example, a player walking aggressively toward a virtual juvenile, finger pointed at him while saying 'What do you think you're doing?' may, depending on the context of the interaction, be interpreted as an aggressive stance. In turn, this would influence the juvenile's response based on his current stance and personality. To implement this approach, future research should investigate the incorporation of social signal processing through computer vision and speech recognition (Vinciarelli, Pantic & Bourlard, 2009).

1 See <http://www.re-lion.com>.

With our Wizard-of-Oz setup of LOITER-VR, we performed several demonstrations, receiving valuable comments to further develop the system. One point of concern was the limited variability in character responses, as became obvious after multiple playthroughs of the same scenario. In our setup, the virtual juveniles had a small set of possible responses, which diminished the replayability of the game. As discussed above, the automated creation of a sufficiently large set of actions is still an open challenge in the field of interactive storytelling. Future work may also profit from expanding on the approach used for the VST and by putting the concepts concerning meta-thinking of virtual agents to good use.

An additional line of research lies in comparing differences between the interactions with LOITER-VR and the 2D LOITER games, for example, to determine learning effects. The improved graphical fidelity of the former, combined with its more immersive experience as compared to the latter, may provide a more realistic experience of difficult social conflicts (Seidel & Chatelier, 2013). This immersion in virtual reality may go as far as producing a ‘shock effect’ of encountering a non-cooperative or even aggressive civilian, which may in turn result in larger learning gains. Moreover, the increased agency players have in LOITER-VR as compared to the 2D LOITER games may also affect their experience, as their available actions are not pre-constructed, but they have to come up with them themselves in real time.

To go beyond a simulation of difficult social situations police officers have to deal with, LOITER-VR could benefit from the use of meta-techniques to improve its learning effects. An open line of enquiry is whether meta-techniques as discussed in Chapter 6 would work in a VR setting, especially when the graphical fidelity is improved. Showing thought bubbles in a highly realistic environment may break players’ immersion, as such objects do not exist in the real world. In the 2D LOITER games, this was less of an intrusion, as the thought bubbles fit the overall style of the game. In the VR environment, this may not be the case. Therefore, special care has to be taken of implementing this meta-technique; future research should point out the effects of including such elements. An alternative may be to use voice-overs to express the virtual juvenile’s thoughts, as they are also used in other media, such as films. Flashbacks are also often featured in films and may therefore may a good fit for a more realistic visualisation of a story world. However, flashbacks should still be presented explicitly as being flashbacks and not currently transpiring events. Otherwise, they might miss their intended goal of communicating information about something which has already happened. For example, a stereotypical black and white filter can be applied to the virtual world as a cue that the visualised scene happens in the past. This may be an indication which players can relate to, as the same effect is used in other media, such as films.

Another major challenge for interactions in VR in general is the sense of touch. Haptic feedback can be very informative to players to discern objects they are interacting with. An example is a player attempting to comfort a virtual juvenile by patting him on the shoulder. This implies that the player needs to be able to feel the juvenile’s shoulder, receiving information on where it is located in the virtual environment. Additionally, this touch needs to be recognized by the system and interpreted in a correct manner (see, for example, Jung, Poppe, Poel and Heylen

(2014)). Depending on the strength of the touch and the cognitive state of the juvenile, he may interpret it as, indeed, comforting, or perhaps as a threat.

The LOITER-VR installation was intended as a proof-of-concept to investigate what is possible with current-day technology and parts of the research discussed in this dissertation. The main challenge of making such a system work lies in the integration of all its parts: artificial intelligence for the virtual characters, the virtual environment, the sensing and interpretation of player behaviour, and the incorporation of mechanics to stimulate learning. With VR outgrowing its stages of infancy, future research should scrutinize how we can improve the affordances of this technology, adapting its mechanics to further improve its experience, both for education and entertainment.

9.3 THE NARRATIVE OF THE FUTURE

The novel of the future provides the experiences we desire and the lessons we require. Like the stories of yore, these experiences should provide current-day morals which fit well with society. Moreover, they should adapt themselves over time, not clinging to old ideals, but grow together with our culture. Interactive media such as choose-your-own-adventure novels and games already entertain and inform us, but we can do more. Murray's view of the Holodeck as the ultimate interactive story is already more than just a dot on the horizon with the rise of virtual reality (Murray, 1997). Stephenson's vision of an educational, adaptive storybook which grows together with a child over time does not seem as far-fetched any more as it did twenty years ago (Stephenson, 1996). I would not be surprised if sooner, rather than later, the following might come to pass.

She feels like she didn't really have a choice to begin with. The cards had been dealt and this could have been the only possible outcome. Her gaze travels upwards, away from the miserable sight of the grotesque, lupine shape, strangely contorted. The poison had done its job alright.

'What's done is done,' the boy informs her. As if she didn't know. 'Though I cannot help but think what if...'

She offers him a soft chuckle and throws her gaze to the left. 'Why wonder?' A rift appears in the air next to her, a grey mirror into the past. The boy steps up to her. Together, they stare at their five-minute-old selves through the rift. They are facing the large, hunched-over wolf, moments before it were to have its last meal.

'It's good we brought the cake, right?' The girl gives the boy a sideways look, then stares at the cake her former self was presenting to the wolf. The boy extends his arm to the rift and taps its side, bringing up a selection panel. His fingers hover over a cloud-shaped icon which would show the wolf's thoughts. But before he can touch it, the girl grabs the boy by his arm and drags him with her into the portal. It only takes one blink to find themselves facing the wolf, once again.

‘Drama is conflict, right?’ she mockingly asks him, dropping the cake from her hands. It lands in the forest soil with a soft thud. The boy’s eyes grow as wide as the girl’s grin.

Squinting at the duo, the wolf licks its teeth, its stomach grumbling. ‘Isn’t that a pity?’ it inquires. ‘Now I have to be content with some other, equally sweet, treat.’

Composing himself, the boy shakes his head. ‘Why settle for just the two of us?’ His smile does not betray his true intentions, a ruse shaping in his mind. ‘Wouldn’t you rather visit our grandmother? She can whip you up as many cakes as you can eat!’ Making a gesture in the air with two fingers, the boy freezes time momentarily and faces the girl, the wolf standing dead still. She quickly catches on.

‘The leafs, just up ahead, right?’ A wink and a nod from the boy is all she needs to grin again. With a flick of her wrist, she undoes the pause and the wolf audibly starts heaving its mighty chest again, drooling at the thought of many more cakes.

It growls: ‘No tricks, you two...’

The children nod solemnly, their eyes twinkling up at the wolf’s yellowish gaze. Treading carefully up the path, they cross a thick carpet of leafs. With the wolf right behind them, the boy and the girl suddenly leap forward. The wolf only has a brief moment to attempt to lunge at them. It fails. The leafs give way under the wolf’s weight and it falls into a deep, dark pit. Its gaze furious, breath fuming, it roars and claws at the children, but they are safely out of reach.

‘Have fun looking at the birds in the sky, Mr. Wolf!’ the girl snickers.

The boy salutes the wolf and gives the girl a high-five: ‘I think we did pretty well!’

‘Yeah, the cake was too obvious, this trap was way more fun,’ the girl replies. ‘I’m curious how we will run into the wolf again next time!’

Smiling, the girl opens and closes her hand two times. Their surroundings fade and the forest dissolves, leaking colour until a living room becomes visible through the haze. The leaves and the pit morph into a thick carpet, the trees turn out to be large closets. Finally, the simulation halts and reality pops in.

Two children, a girl and a boy, are standing on opposite sides of a table.

BIBLIOGRAPHY

- 2K Boston & 2K Australia. (2007). Bioshock. PC, PlayStation 3, Xbox 360. (Cited on p. 19).
- Abt, C. C. (1987). *Serious games*. Lanham, MD: University Press of America. (Cited on p. 8).
- Adams, E. (2008, May 14). Difficulty modes and dynamic difficulty adjustment. Retrieved February 22, 2017, from http://www.designersnotebook.com/Columns/095_Difficulty_Modes/095_difficulty_modes.htm. (Cited on p. 129)
- Albert, A., Hallowell, M. R., Kleiner, B., Chen, A. & Golparvar-Fard, M. (2014). Enhancing construction hazard recognition with high-fidelity augmented virtuality. *Journal of Construction Engineering and Management*, 140(7). (Cited on p. 8).
- Alexander, A. L., Brunyé, T., Sidman, J. & Weil, S. A. (2005). From gaming to training: a review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games. *DARWARS Training Impact Group*, 5, 1–14. (Cited on p. 97).
- Allain, S. & Szilas, N. (2012). Exploration de la métalepse dans les "serious games" narratifs. *Revue des sciences et techniques de l'information et de la communication pour l'éducation et la formation*, 19, 1–14. (Cited on p. 108).
- Alofs, T., Theune, M. & Swartjes, I. (2015). A tabletop interactive storytelling system: designing for social interaction. *International Journal of Arts and Technology*, 8(3), 188–211. (Cited on pp. 8, 33–35, 37, 39, 40, 57, 61, 157, 191).
- Alves, A., Lopes, R., Matos, P., Velho, L. & Silva, D. (2010). Reactoon: Storytelling in a Tangible Environment. In *2010 Third IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning* (pp. 161–165). (Cited on p. 39).
- Amaro, S., Viggiano, A. [Alessandro], Di Costanzo, A., Madeo, I., Viggiano, A. [Andrea], Baccari, M. E., ... Deepak, S. et al. (2006). Kaledo, a new educational board-game, gives nutritional rudiments and encourages healthy eating in children: a pilot cluster randomized trial. *European Journal of Pediatrics*, 165(9), 630–635. (Cited on p. 9).
- Anderson, L. W., Krathwohl, D. R. & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Boston, MA: Allyn & Bacon. (Cited on pp. 94, 95).
- Arinbjarnar, M., Barber, H. & Kudenko, D. (2009). A critical review of interactive drama systems. In *AISB 2009 Symposium. AI & Games*. (Cited on p. 7).
- Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., de Freitas, S., Louchart, S., ... De Gloria, A. (2014). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology*. (Cited on pp. 10, 106).
- Aylett, R. (1999). Narrative in Virtual Environments – Towards Emergent Narrative. *Proceedings of the AAAI fall symposium on narrative intelligence*, 83–86. (Cited on p. 7).

- Aylett, R., Dias, J. & Paiva, A. (2006). An affectively-driven planner for synthetic characters. In *Proceedings of the International Conference on Autonomous Planning and Scheduling* (pp. 2–10). (Cited on pp. 24, 31, 66, 158).
- Aylett, R., Louchart, S., Dias, J., Paiva, A. & Vala, M. (2005). FearNot! An experiment in emergent narrative. In *Proceedings of International Conference on Intelligent Virtual Agents (IVA '05)* (pp. 305–316). (Cited on pp. 38, 39, 108).
- Aylett, R., Louchart, S., Tychsen, A., Hitchens, M., Figueiredo, R. & Mata, C. (2008). Managing emergent character-based narrative. In *Proceedings of the 2nd international conference on intelligent technologies for interactive entertainment* (pp. 1–8). (Cited on pp. 25, 78, 88).
- Bailey, C. & Katchabaw, M. J. (2005). An Experimental Testbed to Enable Auto-Dynamic Difficulty in Modern Video Games. *2005 GameOn North America Conference*, 18–22. (Cited on pp. 126, 129).
- Bal, M. (2009). *Narratology: introduction to the theory of narrative*. Toronto: University of Toronto Press. (Cited on p. 29).
- Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD research*, 1(1). (Cited on p. 12).
- Bates, J., Loyall, A. B. & Reilly, W. S. (1992). An architecture for action, emotion, and social behavior. In *Artificial social systems* (pp. 55–68). (Cited on pp. 66, 72, 156).
- Bateson, G. (1955). A theory of play and fantasy; a report on theoretical aspects of the project of study of the role of the paradoxes of abstraction in communication. *Psychiatric Research Reports*, (2), 39–51. (Cited on p. 6).
- Bateson, G. (1972). A theory of play and fantasy. In G. Bateson (Ed.), *Steps to an ecology of mind* (pp. 177–193). Reprinted in Salen, K. and Zimmerman, E. (eds.), *The Game Design Reader: A Rules of Play Anthology*, pp. 314–318. MIT Press (2006). New York: Chandler. (Cited on pp. 5, 53).
- Belarbi, S., Bergström, K., Ebbenhøj, S. L., Hansen, E. E., Fatland, E., Giæver, O. P., ... Westlund, A. (2010). *Nordic larp* (J. Stenros & M. Montola, Eds.). (Cited on p. 109).
- Bellotti, F., Berta, R., De Gloria, A. & Primavera, L. (2009). Adaptive Experience Engine for Serious Games. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(4), 264–280. (Cited on p. 125).
- Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouvet, D., ... Wellekens, C. (2007). Automatic speech recognition and speech variability: A review. *Speech Communication*, 49(10–11), 763–786. (Cited on p. 166).
- Bethesda Game Studios. (2006). *The Elder Scrolls IV: Oblivion*. PC, Xbox 360, PlayStation 3. (Cited on p. 129).
- BioWare. (2012). *Mass Effect Trilogy*. PC, Xbox 360, PlayStation 3. (Cited on p. 128).
- Black Isle Studios. (1999). *Planescape: Torment*. PC. (Cited on p. 125).
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. & Krathwohl, D. R. (1956). *Taxonomy of educational objectives. Vol. 1: Cognitive domain*. New York, NY: David McKay Company. (Cited on p. 93).
- Bogost, I. (2007). *Persuasive games: the expressive power of videogames*. Cambridge, MA: MIT Press. (Cited on p. 8).

- Bosse, T., Gerritsen, C., de Man, J. & Tolmeijer, S. (2015). Adaptive Training for Aggression de-Escalation. In *Artificial Life and Intelligent Agents* (pp. 80–93). Communications in Computer and Information Science. (Cited on p. 130).
- Bosse, T. & Provoost, S. (2015). Integrating Conversation Trees and Cognitive Models within an ECA for Aggression De-escalation Training. In *Proceedings of Principles and Practice of Multi-Agent Systems '15* (pp. 650–659). (Cited on pp. 87, 130).
- Breuer, J. J. & Bente, G. (2010). Why so serious? On the relation of serious games and learning. *Eludamos. Journal for Computer Game Culture*, 4(1), 7–24. (Cited on pp. 8, 126).
- Brockmyer, J. H., Fox, C. M., Curtiss, K. a., Mcbroom, E., Burkhart, K. M. & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire : A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4), 624–634. (Cited on p. 143).
- Broekens, J., Harbers, M., Brinkman, W.-P., Jonker, C. M., van den Bosch, K. & Meyer, J.-J. C. (2012). Virtual reality negotiation training increases negotiation knowledge and skill. In *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '12)* (pp. 218–230). (Cited on pp. 108, 112).
- Brown, A. D., Stacey, P. & Nandhakumar, J. (2008). Making sense of sensemaking narratives. *Human Relations*, 61(8), 1035–1062. (Cited on p. 5).
- Brown, P. & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge, MA: Cambridge University Press. (Cited on p. 91).
- Bruijnes, M. (2016). *Believable Suspect Agents: Response and Interpersonal Style Selection for an Artificial Suspect* (PhD thesis, University of Twente). (Cited on pp. 89, 90).
- Bruijnes, M., Linssen, J. M., op den Akker, R., Theune, M., Wapperom, S., Broekema, C. & Heylen, D. K. J. (2015). Social Behaviour in Police Interviews: Relating Data to Theories. In *Conflict and multimodal communication* (pp. 317–347). (Cited on pp. 89, 91, 92, 133).
- Bruijnes, M., op den Akker, R., Hartholt, A. & Heylen, D. K. J. (2015). Virtual Suspect William. In W.-P. Brinkman, J. Broekens & D. Heylen (Eds.), *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '15)* (Vol. 9238, pp. 67–76). (Cited on p. 90).
- Bruijnes, M., op den Akker, R., Spitters, S., Sanders, M. & Fu, Q. (2015). The recognition of acted interpersonal stance in police interrogations and the influence of actor proficiency. *Journal on Multimodal User Interfaces*, 9(4), 353–376. (Cited on p. 90).
- Bunt, H. (2009). The DIT++ taxonomy for functional dialogue markup. In *Proceedings of the AAMAS 2009 Workshop 'Towards a Standard Markup Language for Embodied Dialogue Acts'*. (Cited on pp. 166, 168).
- Caillois, R. (2001). *Man, Play and Games*. Champaign, IL: University of Illinois Press. (Cited on pp. 9, 155, 156).
- Campbell, J. (1949). *The Hero with a Thousand Faces*. Novato, CA: New World Library. (Cited on pp. 5, 156).
- Cavazza, M., Charles, F. & Mead, S. (2002). Character-based interactive storytelling. *IEEE Intelligent Systems*, 17(4), 17–24. (Cited on p. 67).

- Cavazza, M., Lugin, J., Pizzi, D. & Charles, F. (2007). Madame Bovary on the Holodeck: Immersive interactive storytelling. In *Proceedings of the 15th ACM International Conference on Multimedia* (pp. 651–660). (Cited on pp. 22, 23).
- Chang, H. M. & Soo, V. W. (2009). Planning-based narrative generation in simulated game universes. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(3), 200–213. (Cited on pp. 67, 79).
- Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., ... Popovi, Z. et al. (2010). Predicting protein structures with a multiplayer online game. *Nature*, 466(7307), 756–760. (Cited on p. 8).
- Core, M., Traum, D., Lane, H. & Swartout, W. (2006). Teaching negotiation skills through practice and reflection with virtual humans. *Simulation*, 82(11), 685–701. (Cited on p. 108).
- Csikszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: HarperCollins. (Cited on pp. 126, 127, 151).
- Davis-Mendelow, S., Baier, J. A. & McIlraith, S. A. (2012). Making reasonable assumptions to plan with incomplete information: Abridged report. *AAAI Workshop – Technical Report, WS-12-12*, 16–22. (Cited on p. 67).
- de Kort, Y., IJsselstein, W. A. & Poels, K. (2007). Digital Games as Social Presence Technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). In *Proceedings of the 10th Annual International Workshop on Presence* (pp. 195–203). (Cited on p. 143).
- Dever, J. & Chalk, G. (1984). *Lone Wolf: Flight From the Dark*. New York, NY: Berkley Books. (Cited on p. 6).
- Djaouti, D., Alvarez, J. & Jessel, J.-P. (2011). Classifying serious games: The G/P/S model. *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches*, 118–136. (Cited on pp. 8, 9).
- Djaouti, D., Alvarez, J., Jessel, J.-p. & Rampnoux, O. (2011). Origins of Serious Games. In M. Ma, A. Oikonomou & L. C. Jain (Eds.), *Serious games and edutainment applications* (pp. 25–43). (Cited on p. 8).
- Donker, A. & Reitsma, P. (2004). Usability testing with young children. In *Proceedings of the 2004 conference on Interaction design and children building a community* (pp. 43–48). (Cited on p. 52).
- Duke, R. D. & Kriz, W. (2014). *Back to the future of gaming*. Bielefeld: W. Bertelsmann Verlag. (Cited on p. 83).
- Dunwell, I., Freitas, S. D. & Jarvis, S. (2011). Four-dimensional consideration of feedback in serious games. *Digital Games and Learning*, 42–62. (Cited on p. 106).
- Egan, G. (1995). Into Darkness. In *Axiomatic*. London: Victor Gollancz Ltd. (Cited on p. 165).
- Epic MegaGames & Digital Extremes. (1998). Unreal. PC. (Cited on p. 129).
- Erikson, S. (2009). *House of Chains*. London: Transworld Publishers. (Cited on p. 83).
- Faas, S. (2002). *Virtual Storyteller – An approach to computational story telling* (MSc thesis, University of Twente). (Cited on pp. 28, 29, 165).
- Fine, G. A. (2002). *Shared fantasy: role playing games as social worlds*. Chicago, IL: University of Chicago Press. (Cited on pp. 53–55, 62, 158).

- Fisher, W. R. (1989). *Human communication as narration: toward a philosophy of reason, value, and action*. Columbia, South Carolina: University of South Carolina Press. (Cited on pp. 5, 38).
- Flaubert, G. (2004). *Madame Bovary* (M. Mauldon, Trans.). Oxford, UK: Oxford University Press. (Cited on p. 22).
- Freytag, G. (1863). *Die Technik des Dramas*. Stuttgart: S. Hirzel Verlag. (Cited on p. 26).
- Frissen, V., de Lange, S., de Mul, J. & Raessens, J. (2015). *Playful Identities: The Ludification of Digital Media Cultures*. Amsterdam: Amsterdam University Press. (Cited on p. 11).
- Göbel, S., Mehm, F., Radke, S. & Steinmetz, R. (2009). 80days: Adaptive digital storytelling for digital educational games. In *Proceedings of the 2nd International Workshop on Story-Telling and Educational Games (STEG '09)*. (Cited on p. 130).
- Goffman, E. (1959). *The presentation of self in everyday life*. New York, NY: Garden City. (Cited on pp. 91, 139).
- Goffman, E. (1974). *Frame analysis: an essay on the organization of experience*. Cambridge, MA: Harvard University Press. (Cited on p. 53).
- Gordon, A. & Lent, M. V. (2002). Virtual Humans as Participants vs. Virtual Humans as Actors. In *Proceedings of the AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment* (pp. 34–38). (Cited on p. 68).
- Graesser, A. C., Olde, B. & Klettke, B. (2002). How does the mind construct and represent stories? *Narrative Impact: Social and Cognitive Foundations*, 6400(901), 229–262. (Cited on p. 5).
- Gratch, J. & Marsella, S. (2001). Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In *Proceedings of the fifth international conference on Autonomous Agents* (pp. 278–285). (Cited on p. 84).
- Grimm, J. & Grimm, W. (2008). Little red-cap (E. Taylor & M. Edwardes, Trans.). In, *Fairy tales*. Salt Lake City, UT: Project Gutenberg. (Cited on pp. 33, 37).
- Gupta, S., Walker, M. & Romano, D. (2007). How rude are you? Evaluating politeness and affect in interaction. In *Proceedings of Affective Computing and Intelligent Interaction (ACII '07)* (pp. 203–217). (Cited on p. 127).
- Gurtman, M. B. (2009). Exploring Personality with the Interpersonal Circumplex. *Social and Personality Psychology Compass*, 3(4), 601–619. (Cited on pp. 90, 134).
- Ha, E. Y., Rowe, J. P., Mott, B. W. & Lester, J. C. (2012). Goal Recognition with Markov Logic Networks for Player-Adaptive Games. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence* (pp. 2113–2119). (Cited on p. 166).
- Habonneau, N., Richle, U., Szilas, N. & Dumas, J. E. (2012). 3D Simulated Interactive Drama for Teenagers Coping with a Traumatic Brain Injury in a Parent. In *Proceedings of the International Conference on Interactive Digital Storytelling (ICIDS '12)* (pp. 174–182). (Cited on pp. 88, 89).
- Helmes, J., Cao, X., Lindley, S. E. & Sellen, A. (2009). Developing the story: Designing an Interactive Storytelling Application. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '09)* (pp. 49–52). (Cited on p. 39).
- Hill Jr., R., Belanich, J., Lane, H. & Core, M. (2006). Pedagogically structured game-based training: Development of the ELECT BiLAT simulation. In *Proceedings of the 25th Army Science Conference*. (Cited on p. 86).

- Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E. & Singer, D. (2008). *A mandate for playful learning in preschool: applying the scientific evidence*. Oxford, UK: Oxford University Press. (Cited on p. 6).
- Hoek, M. (2014). *Generating Dutch focalized stories about interaction in a serious game* (Master's thesis, University of Twente). (Cited on p. 29).
- Huizinga, J. (1949). *Homo Ludens: A Study of the Play-Element in Culture*. London: Routledge & Kegan Paul Ltd. (Cited on pp. 5, 9, 156).
- id Software. (1993). DOOM. PC. (Cited on p. 128).
- Ijsselstein, W. A., de Kort, Y. A. W. & Poels, K. (n.d.). *The Game Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital games*. (Cited on p. 144).
- Jennett, C., Cox, A. & Cairns, P. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9), 641–661. (Cited on pp. 110, 111).
- Johnson, D. W. & Johnson, R. T. (1990). Social skills for successful group work. *Educational leadership*, 47(4), 29–33. (Cited on p. 83).
- Johnson, W. & Friedland, L. (2010). Integrating cross-cultural decision making skills into military training. *Advances in Cross-Cultural Decision Making*. (Cited on pp. 85, 86).
- Johnstone, K. (2012). *Impro: improvisation and the theatre*. Abingdon-on-Thames: Routledge. (Cited on pp. 7, 31).
- Jung, M. M., Poppe, R., Poel, M. & Heylen, D. K. (2014). Touching the Void – Introducing CoST: Corpus of Social Touch. In *Proceedings of the 16th International Conference on Multimodal Interaction (ICMI '14)* (pp. 120–127). (Cited on p. 169).
- Kaldenbach, H. (2006). *Respect! 99 tips voor het omgaan met jongeren in de straatcultuur*. Amsterdam: Prometheus. (Cited on p. 96).
- Kaldenbach, H. (2008). *Hangjongeren, 99 tips voor buurtbewoners en voorbijgangers*. Amsterdam: Prometheus. (Cited on p. 96).
- Kickmeier-Rust, M. D., Mattheiss, E., Steiner, C. M. & Albert, D. (2011). A psychopedagogical framework for multi-adaptive educational games. *International Journal on Game-Based Learning*, 1(1), 45–58. (Cited on p. 126).
- Kickmeier-Rust, M. D., Schwarz, D., Albert, D., Verpoorten, D., Castaigne, J. L. & Bopp, M. (2006). The ELEKTRA project: Towards a new learning experience. In *M3: Interdisciplinary aspects on digital media & education* (Vol. 3, pp. 19–48). (Cited on p. 130).
- Kickmeier-Rust, M. D., Steiner, C. M. & Albert, D. (2011). Apt to adapt: Micro- and macro-level adaptation in educational games. In T. Daradoumis, S. Caballé, A. A. Juan & F. Xhafa (Eds.), *Technology-enhanced systems and tools for collaborative learning scaffolding* (pp. 221–238). (Cited on p. 130).
- Kiili, K. (2007). Foundation for problem-based gaming. *British Journal of Educational Technology*, 38(3), 394–404. (Cited on pp. 106, 107).
- Kim, J. M., Hill Jr., R. W., Durlach, P. J., Lane, H. C., Forbell, E., Mark Core, ... Hart, J. (2009). BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education*, 19, 289–308. (Cited on pp. 86, 108).

- Kirriemuir, J. & McFarlane, A. (2004). *Literature review in games and learning*. A NESTA Futurelab Research report – report 8. (Cited on p. 6).
- Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. (Cited on p. 106).
- Koops, M. & Hoevenaars, M. (2012). Conceptual Change During a Serious Game: Using a Lemniscate Model to Compare Strategies in a Physics Game. *Simulation & Gaming*, 44(4), 544–561. (Cited on pp. 10, 11, 107, 111, 124).
- Krathwohl, D. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212–218. (Cited on pp. 94, 95).
- Kulhavy, R. W. & Wager, W. (1993). Feedback in programmed instruction: historical context and implications for practice. In J. V. Dempsey & G. C. Sales (Eds.), *Interactive instruction and feedback*. (Cited on p. 106).
- Laamarti, F., Eid, M. & El Saddik, A. (2014). An Overview of Serious Games. *International Journal of Computer Games Technology*, 1–15. (Cited on p. 9).
- Laffey, J., Schmidt, M., Galyen, K. & Stichter, J. (2012). Smart 3D collaborative virtual learning environments: A preliminary framework. *Journal of Ambient Intelligence and Smart Environments*, 4(1). (Cited on p. 127).
- Lane, H., Core, M. & Lent, M. V. (2005). *Explainable artificial intelligence for training and tutoring*. University of California, Institute for Creative Technologies. (Cited on p. 107).
- Leary, T. (1957). *Interpersonal diagnosis of personality: functional theory and methodology for personality evaluation*. New York: Ronald Press. (Cited on pp. 86, 90, 94, 98, 132, 134, 152, 160, 168).
- Lem, S. (2011). *Solaris* (B. Johnston, Trans.). Los Angeles, CA: Premier Digital Publishing. (Cited on p. 105).
- Lethbridge, S. & Mildorf, J. (2003). *Basics of English Studies: An introductory course for students of literary studies in English*. Tübingen: University of Tübingen. (Cited on pp. 6, 26).
- Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D. & Palmquist, C. M. (2013). The impact of pretend play on children's development: a review of the evidence. *Psychological bulletin*, 139(1), 1. (Cited on p. 6).
- Linssen, J. M. (2011). *Adaptive Learning in an Educational Game - Adapting Game Complexity to Gameplay Increases Efficiency of Learning* (Master's thesis, Utrecht University, Utrecht). (Cited on p. 129).
- Linssen, J. M. (2012). *A Discussion of Interactive Storytelling Techniques for Use in a Serious Game*. University of Twente, Enschede: Centre for Telematics and Information Technology, University of Twente. (Cited on pp. 19, 83).
- Linssen, J. M. (2014). Meta-techniques for a Social Awareness Learning Game. In *European Conference on Games Based Learning* (pp. 697–704). (Cited on pp. 94, 105).
- Linssen, J. M. & de Groot, T. F. (2013). AGENT: Awareness Game Environment for Natural Training. In *Proceedings of the International Conference on the Foundations of Digital Games '13*. (Cited on p. 167).

- Linssen, J. M., de Groot, T. F., Theune, M. & Bruijnes, M. (2014). Beyond simulations: serious games for training interpersonal skills in law enforcement. In *Proceedings of the 10th annual meeting of the European Social Simulation Association (ESSA 2014)*. (Cited on p. 97).
- Linssen, J. M., de Groot, T. F., Theune, M. & Heylen, D. K. J. (2015). LOITER-TB: Thought Bubbles that Give Feedback on Virtual Agents' Experiences. In *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '15)* (pp. 283–286). (Cited on pp. 83, 96).
- Linssen, J. M. & Theune, M. (2012). Using out-of-character reasoning to combine storytelling and education in a serious game. In *CEUR Workshop Proceedings* (Vol. 898, pp. 27–30). (Cited on p. 125).
- Linssen, J. M., Theune, M. & de Groot, T. F. (2013). What Is at Play? Meta-techniques in Serious Games and Their Effects on Social Believability and Learning. In *Proceedings of the Social Believability in Games Workshop '13*. (Cited on p. 105).
- Linssen, J. M., Theune, M., de Groot, T. F. & Heylen, D. K. J. (2015). Improving Social Awareness Through Thought Bubbles and Flashbacks of Virtual Characters. In *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '15)* (pp. 250–259). (Cited on p. 105).
- Lohse, K. R., Hilderman, C. G. E., Cheung, K. L., Tatla, S. & Van der Loos, H. F. (2014). Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. *PloS one*, 9(3), e93318. (Cited on p. 8).
- Louchart, S. & Aylett, R. (2003). Solving the Narrative Paradox in VEs - Lessons from RPGs. In *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '03)* (pp. 244–248). (Cited on pp. 49, 52).
- Low, J. & Durkin, K. (1998). Structure and causal connections in children's on-line television narratives: What develops? *Cognitive Development*, 13(2), 201–225. (Cited on pp. 38, 43, 46, 61).
- Lugrin, J., Cavazza, M., Pizzi, D., Vogt, T. & André, E. (2010). Exploring the Usability of Immersive Interactive Storytelling. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology* (pp. 103–110). (Cited on p. 23).
- Lukka, L. (2011). The dual-faceted role. In T. D. Henriksen, C. Bierlich, K. F. Hansen & V. Kølle (Eds.), *Think LARP: Academic Writings from KP2011* (pp. 152–170). Copenhagen, DK: Rollespilsakademiet. (Cited on p. 111).
- Lynch, J. S. & van den Broek, P. (2007). Understanding the glue of narrative structure: Children's on- and off-line inferences about characters' goals. *Cognitive Development*, 22(3), 323–340. (Cited on p. 52).
- Marsella, S., Johnson, W. L. & LaBore, C. (2000). Interactive pedagogical drama. *Proceedings of the fourth international conference on Autonomous agents - AGENTS '00*, 301–308. (Cited on p. 110).
- Marsella, S., Pynadath, D. & Read, S. (2004). PsychSim: Agent-based modeling of social interactions and influence. In *Proceedings of the International Conference on Cognitive Modeling '04* (pp. 243–248). (Cited on p. 85).

- Mason, S. (2013). On Games and Links: Extending the Vocabulary of Agency and Immersion in Interactive Narratives. In *Proceedings of the International Conference on Interactive Digital Storytelling '13* (pp. 25–34). (Cited on p. 133).
- Mateas, M. & Stern, A. (2002). Towards Integrating Plot and Character for Interactive Drama. In K. Dautenhahn, A. Bond, L. Cañamero & B. Edmonds (Eds.), *Socially intelligent agents* (pp. 221–228). (Cited on p. 7).
- Mateas, M. & Stern, A. (2003). Façade: An Experiment in Building a Fully-Realized Interactive Drama. In *Proceedings of the Game Developers Conference '03* (Vol. 2). (Cited on pp. 7, 25, 26, 166).
- Mateas, M. & Stern, A. (2005a). Procedural authorship: A case-study of the interactive drama Façade. In *Digital Arts and Culture (DAC)*. (Cited on p. 27).
- Mateas, M. & Stern, A. (2005b). Structuring content in the Façade interactive drama architecture. *Proceedings of the First Artificial Intelligence and Interactive Digital Entertainment Conference, June 1-5, 2005, Marina del Rey, California, USA*, 3, 93–98. (Cited on p. 52).
- McClure, D. (2009, April 12). Adaptive difficulty. Retrieved February 22, 2017, from http://www.gamecareerguide.com/features/805/adaptive_.php. (Cited on p. 129)
- Meehan, J. R. (1977). TALE-SPIN, an interactive program that writes stories. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence (IJCAI '77)* (pp. 91–98). (Cited on p. 66).
- Meier, C. R., DiPerna, J. C. & Oster, M. M. (2006). Importance of social skills in the elementary grades. *Education and Treatment of Children*, 29(3), 409–419. (Cited on p. 83).
- Morgeson, F. P., Reider, M. H. & Campion, M. A. (2005). Selecting individuals in team settings: the importance of social skills, personality characteristics, and teamwork knowledge. *Personnel Psychology*, 58(3), 583–611. (Cited on p. 83).
- Mory, E. H. (2004). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745–783). (Cited on p. 106).
- Mott, B. W. & Lester, J. C. (2006). U-director: a decision-theoretic narrative planning architecture for storytelling environments. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '06)* (pp. 977–984). (Cited on p. 21).
- Murray, J. H. (1997). *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. New York, NY: Free Press. (Cited on pp. 6, 7, 170).
- Nestel, D. & Tierney, T. (2007). Role-play for medical students learning about communication: guidelines for maximising benefits. *BMC Medical Education*, 7(1), 3. (Cited on p. 9).
- Niles, J. D. (1999). *Homo narrans: the poetics and anthropology of oral literature*. Philadelphia, PA: University of Pennsylvania Press. (Cited on p. 5).
- Nintendo. (1992). Mario Kart. Super Nintendo Entertainment System. (Cited on p. 128).
- Ortony, A., Clore, G. L. & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge, UK: Cambridge University Press. (Cited on p. 24).

- Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. & Fuller, T. (2002). User-centered Design in Games. In J. Jacko & A. Sears (Eds.), *Handbook for human-computer interaction in interactive systems* (pp. 883–906). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (Cited on p. 128).
- Paiva, A., Machado, I. & Prada, R. (2001). The child behind the character. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 31(5), 361–368. (Cited on p. 39).
- Paris, C., Swartout, W. R. & Mann, W. C. (2013). *Natural language generation in artificial intelligence and computational linguistics*. Berlin: Springer Science & Business Media. (Cited on p. 166).
- Peeters, M. M., van den Bosch, K., Meyer, J.-J. C. & Neerincx, M. A. (2014). The design and effect of automated directions during scenario-based training. *Computers & Education*, 70, 173–183. (Cited on pp. 131, 132).
- Peinado, F. & Gervás, P. (2006). Evaluation of automatic generation of basic stories. *New Generation Computing*, 24(3), 289–302. (Cited on p. 79).
- Peirce, N., Conlan, O. & Wade, V. (2008). Adaptive Educational Games: Providing Non-invasive Personalised Learning Experiences. In *Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning* (pp. 28–35). (Cited on pp. 129, 130).
- Pereira, G., Brisson, A., Prada, R., Paiva, A., Bellotti, F., Kravcik, M. & Klamma, R. (2012). Serious Games for Personal and Social Learning & Ethics: Status and Trends. *Procedia Computer Science*, 15, 53–65. (Cited on p. 83).
- Pizzi, D. & Cavazza, M. (2007). Affective Storytelling based on Characters' Feelings. In *Proceedings of the AAAI Fall Symposium on Intelligent Narrative Technologies '07* (pp. 110–117). (Cited on pp. 22, 23).
- Prada, R., Paiva, A., Machado, I. & Gouveia, C. (2002). 'You Cannot Use My Broom! I'm the Witch, You're the Prince': Collaboration in a Virtual Dramatic Game. *Proceedings of Intelligent Tutoring Systems, 6th International Conference (ITS '02)*, 913–922. (Cited on pp. 38, 39).
- Quintana, C., Reiser, B. J., Davis, E. a., Krajcik, J., Fretz, E., Duncan, R. G., ... Soloway, E. (2009). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 13(3), 37–41. (Cited on p. 127).
- Ravenet, B., Ochs, M. & Pelachaud, C. (2013). From a User-created Corpus of Virtual Agent's Non-verbal Behavior to a Computational Model of Interpersonal Attitudes. In *Intelligent virtual agents* (pp. 263–274). (Cited on pp. 98, 112).
- Reese, E., Haden, C. a., Baker-Ward, L., Bauer, P., Fivush, R. & Ornstein, P. A. (2011). Coherence of Personal Narratives Across the Lifespan: A Multidimensional Model and Coding Method. *Journal of Cognition and Development*, 12(4), 424–462. (Cited on p. 38).
- Remedy Entertainment. (2001). *Max Payne*. PC. (Cited on pp. 128, 155).
- Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D. & Swartout, W. (2002). Toward a new generation of virtual humans for interactive experiences. *Intelligent Systems, IEEE*, 17(4), 32–38. (Cited on pp. 84, 85).
- Ricoeur, P. (1983). *Time and narrative* (K. McLaughlin & D. Pellauer, Trans.). Chicago, IL: University of Chicago Press. (Cited on p. 5).

- Riedl, M. O., Saretto, C. J. & Young, R. M. (2003). Managing interaction between users and agents in a multi-agent storytelling environment. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '03)* (pp. 741–748). (Cited on p. 20).
- Riedl, M. O. & Stern, A. (2006a). Believable agents and intelligent story adaptation for interactive storytelling. In *Proceedings of the International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE '06)* (pp. 1–12). (Cited on p. 27).
- Riedl, M. O. & Stern, A. (2006b). Failing Believably: Toward Drama Management with Autonomous Actors in Interactive Narratives. In *Proceedings of the International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE '06)* (pp. 195–206). (Cited on p. 28).
- Riedl, M. O. & Young, R. M. (2003). Character-Focused Narrative Generation for Execution in Virtual Worlds. In *Proceedings of the International Conference on Virtual Storytelling (VS '03)* (pp. 47–56). (Cited on p. 67).
- Riedl, M. O. & Young, R. M. (2005). An objective character believability evaluation procedure for multi-agent story generation systems. In *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '05)*. (Cited on p. 79).
- Riedl, M. O. & Young, R. M. (2014). The Importance of Narrative as an Affective Instructional Strategy. In R. A. Sottolare, A. C. Graesser, X. Hu & B. S. Goldberg (Eds.), *Design recommendations for intelligent tutoring systems* (Vol. 2, pp. 57–70). (Cited on p. 127).
- Roberts, D. & Isbell, C. (2008). A survey and qualitative analysis of recent advances in drama management. *International Transactions on Systems Science and Applications, Special Issue on Agent Based Systems for Human Learning*, 4(2). (Cited on p. 7).
- Robertson, J. (2000). *The Effectiveness of a Virtual Role-play Environment as a Preparation Activity for Story Writing* (PhD thesis, University of Edinburgh). (Cited on p. 39).
- Robertson, J. & Despa, J. O. (2006). Ghostwriter: Educational Drama and Presence in a Virtual Environment. *Journal of Computer-Mediated Communication*, 8(1). (Cited on p. 39).
- Roth, C. & Koenitz, H. (2016). Evaluating the User Experience of Interactive Digital Narrative. In *Proceedings of the 1st International Workshop on Multimedia Alternate Realities '16* (pp. 31–36). (Cited on p. 79).
- Rouckhout, D. & Schacht, R. (2000). Ontwikkeling van een Nederlandstalig interpersoonlijk circumplex. *Diagnostiekwijzer*, 4, 96–118. (Cited on p. 112).
- Ryan, J. O., Summerville, A., Mateas, M. & Wardrip-fruin, N. (2015). Toward Characters Who Observe, Tell, Misremember, and Lie. In *Proceedings of the 2nd Workshop on Experimental AI in Games '15* (pp. 56–62). (Cited on p. 79).
- Salen, K. & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge, MA: MIT Press. (Cited on pp. 5, 9, 155).
- Sawyer, R. K. (2003). Levels of analysis in pretend play discourse: Metacommunication in conversational routines. *Play and Culture Studies*, 5, 137–158. (Cited on pp. 48, 49, 51, 53, 54, 61, 62, 158).

- Searle, J. R. (1969). *Speech acts: an essay in the philosophy of language*. Cambridge, MA: Cambridge University Press. (Cited on p. 93).
- Seidel, R. J. & Chatelier, P. R. (2013). *Virtual reality, training's future? Perspectives on virtual reality and related emerging technologies*. Berlin: Springer Science & Business Media. (Cited on p. 169).
- Shapiro, L. R. & Hudson, J. a. (1991). Tell me a make-believe story: Coherence and cohesion in young children's picture-elicited narratives. *Developmental Psychology*, 27(6), 960–974. (Cited on pp. 38, 43, 61).
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. (Cited on p. 106).
- Si, M., Marsella, S. & Pynadath, D. (2005). THESPIAN: An architecture for interactive pedagogical drama. In *Proceeding of the 2005 Conference on Artificial Intelligence in Education* (pp. 595–602). (Cited on p. 85).
- Sindlar, M. P., Dastani, M. M., Dignum, F. & Meyer, J.-J. C. (2010). Explaining and Predicting the Behavior of BDI-Based Agents in Role-Playing Games. In *Proceedings of the 7th International Workshop on Declarative Agent Languages and Technologies '10* (pp. 174–191). (Cited on p. 79).
- Slabbers, N. (2006). *Narration for virtual storytelling* (MSc thesis, University of Twente). (Cited on p. 29).
- Smelt, P., Habchi, R., Dekker, D. & van der Wal, W. (2014). *Straatcultuur. Effectieve aanpak van jeugdgroepen*. Duivendrecht: Uitgeverij SWP. (Cited on p. 139).
- Stephenson, N. (1996). *The Diamond Age: Or, A Young Lady's Illustrated Primer*. London: Penguin. (Cited on pp. 3, 170).
- Stone, C. A. (1998). The Metaphor of Scaffolding: Its Utility for the Field of Learning Disabilities. *Journal of Learning Disabilities*, 31(4), 344–364. (Cited on p. 127).
- Stromberg, P. G. (2000). The “I” of enthrallment. *Ethos*, 27(4), 490–504. (Cited on p. 59).
- Swartjes, I. M. T. (2010). *Whose story is it anyway? How improv informs agency and authorship of emergent narrative* (Ph.D. Thesis, University of Twente). (Cited on pp. 8, 19, 29–33, 156, 165).
- Swartjes, I. M. T., Kruizinga, E. & Theune, M. (2008). Let's Pretend I Had a Sword – Late Commitment in Emergent Narrative. In *Proceedings of the International Conference on Interactive Digital Storytelling (ICIDS '15)* (pp. 264–267). (Cited on p. 31).
- Swartjes, I. M. T. & Theune, M. (2006). A Fabula Model for Emergent Narrative. In S. Gobel, R. Malkewitz & I. Iurgel (Eds.), *Proceedings of the Third International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE '06)* (pp. 49–60). (Cited on p. 29).
- Swartjes, I. M. T. & Theune, M. (2009a). Iterative authoring using story generation feedback: debugging or co-creation? In *Proceedings of the International Conference on Interactive Digital Storytelling (ICIDS '09)* (pp. 62–73). (Cited on pp. 52, 62, 158).
- Swartjes, I. M. T. & Theune, M. (2009b). *Late commitment: virtual story characters that can frame their world*. Enschede: Centre for Telematics and Information Technology University of Twente. (Cited on pp. 67, 72).

- Swartout, W. (2010). Lessons Learned from Virtual Humans. *AI Magazine*, 31(1), 9–20. (Cited on pp. 84, 86).
- Szilas, N. (2005). The future of interactive drama. In *Proceedings of the second Australasian Conference on Interactive Entertainment* (pp. 193–199). (Cited on p. 137).
- Szilas, N. (2007). A Computational Model of an Intelligent Narrator for Interactive Narratives. *Applied Artificial Intelligence*, 21(8), 753–801. (Cited on p. 88).
- Tannen, D. (1993). What's in a frame? Surface evidence for underlying expectations. *Framing in discourse*, 14. (Cited on p. 93).
- Telltale Games. (2012). *The Walking Dead*. PC, PlayStation 3, Xbox 360. (Cited on p. 106).
- ten Brinke, H. (2014). *Hide and sneak - Perceptions in The Virtual Storyteller* (Master's thesis, University of Twente). (Cited on p. 65).
- ten Brinke, H., Linssen, J. M. & Theune, M. (2014). Hide and Sneak: Story Generation with Characters that Perceive and Assume. In *Tenth Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE '14)* (pp. 174–180). (Cited on p. 65).
- Theune, M., Alofs, T., Linssen, J. M. & Swartjes, I. (2013). Having one's cake and eating it too: Coherence of children's emergent narratives. In *Proceedings of the 4th Workshop on Computational Models of Narrative (CMN '13)* (Vol. 32, pp. 293–309). (Cited on pp. 37, 39, 41).
- Theune, M., Linssen, J. M. & Alofs, T. (2013). Acting, Playing, or Talking about the Story: An Annotation Scheme for Communication during Interactive Digital Storytelling. In *Proceedings of the 6th International Conference on Interactive Digital Storytelling (ICIDS '13)* (pp. 132–143). (Cited on pp. 37, 57).
- Theune, M., Rensen, S., op den Akker, R., Heylen, D. & Nijholt, A. (2004). Emotional characters for automatic plot creation. In *Proceedings of the 2nd International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE '04)* (pp. 95–100). (Cited on p. 165).
- Thomas, K. (1992). Conflict and conflict management: Reflections and update. *Journal of Organizational Behavior*, 13(3), 265–274. (Cited on p. 93).
- Tickle-Degnen, L. & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, 1(4), 285–293. (Cited on pp. 91, 92).
- Tommassen, P. (2009). *Wacky Pirate Hijinks - Causing and Resolving Conflict between Autonomous Agents in a Virtual Storyworld* (Master thesis, University of Twente). (Cited on p. 73).
- Trabasso, T., Secco, T. & Van den Broek, P. W. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83–111). Hillsdale, NJ: Lawrence Erlbaum Associates. (Cited on pp. 43, 45).
- Trabasso, T., Stein, N. L., Rodkin, P. C., Park Munger, M. & Baughn, C. R. (1992). Knowledge of goals and plans in the on-line narration of events. *Cognitive Development*, 7(2), 133–170. (Cited on pp. 29, 38, 46, 52, 62).
- Traum, D., Marsella, S., Gratch, J., Lee, J., Hartholt, A., Prendinger, H., ... Ishizuka, M. (2008). Multi-party, Multi-issue, Multi-strategy Negotiation for Multi-modal

- Virtual Agents. In *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '08)* (pp. 117–130). (Cited on p. 85).
- Traum, D., Swartout, W., Marsella, S. & Gratch, J. (2005). Fight, Flight, or Negotiate: Believable Strategies for Conversing under Crisis. In *Proceedings of the International Conference on Intelligent Virtual Agents (IVA '05)* (pp. 52–64). (Cited on p. 85).
- Troika Games. (2004). *Vampire: The Masquerade – Bloodlines*. PC. (Cited on p. 65).
- Tseng, J. C. R., Chu, H.-C., Hwang, G.-J. & Tsai, C.-C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers & Education*, 51(2), 776–786. (Cited on p. 125).
- Vaassen, F. & Wauters, J. (2012). deLearyous: Training Interpersonal Communication Skills Using Unconstrained Text Input. In *Proceedings of the European Conference on Games Based Learning (ECGBL '12)* (pp. 505–513). (Cited on pp. 86, 124, 134, 168).
- van der Spek, E. D. (2011). *Experiments in serious game design: a cognitive approach* (PhD thesis, Utrecht University). (Cited on pp. 9, 131).
- van Lehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, 46(4), 197–221. (Cited on pp. 108, 130).
- van Lent, M., Rey, M., Fisher, W. & Mancuso, M. (2004). An Explainable Artificial Intelligence System for Small-unit Tactical Behavior. In *Proceedings of the Sixteenth Conference on Innovative Applications of Artificial Intelligence (IAAI '04)* (pp. 900–907). (Cited on p. 110).
- van Oostendorp, H., van der Spek, E. D. & Linssen, J. (2014, May). Adapting the Complexity Level of a Serious Game to the Proficiency of Players. *EAI Endorsed Transactions on Game-Based Learning*, 1(2), e5. (Cited on p. 131).
- Vannini, N., Enz, S., Sapouna, M., Wolke, D., Watson, S., Woods, S., ... Schneider, W. (2011). 'FearNot!': a computer-based anti-bullying-programme designed to foster peer intervention. *European Journal of Psychology of Education*, 26(1), 21–44. (Cited on p. 108).
- Vermeulen, I. E., Roth, C., Vorderer, P. & Klimmt, C. (2010). Measuring User Responses to Interactive Stories: Towards a Standardized Assessment Tool. In *Proceedings of the Third International Conference on Interactive Digital Storytelling (ICIDS '10)* (Vol. 6432 LNCS, pp. 38–43). (Cited on p. 39).
- Vinciarelli, A., Pantic, M. & Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 27(12), 1743–1759. (Cited on p. 168).
- Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F. & Schroeder, M. (2012). Bridging the Gap between Social Animal and Unsocial Machine: A Survey of Social Signal Processing. *IEEE Transactions on Affective Computing*, 3(1), 69–87. (Cited on p. 110).
- Vygotsky, L. S. (1978a). Interaction between learning and development. In *Mind in society* (pp. 79–91). Cambridge, MA: Harvard University Press. (Cited on pp. 126–128, 151).

- Vygotsky, L. S. (1978b). *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, Ed.). Cambridge, MA: Harvard University Press. (Cited on p. 126).
- Watts, P. (2006). *Blindsight*. New York, NY: Tor Books. (Cited on p. 37).
- Weser, M., Off, D. & Jianwei Zhang. (2010). HTN robot planning in partially observable dynamic environments. In *2010 IEEE International Conference on Robotics and Automation* (pp. 1505–1510). (Cited on p. 67).
- Weyhrauch, P. (1997). *Guiding interactive drama* (Ph.D. Thesis, Carnegie Mellon University). (Cited on p. 25).
- Winer, M., Beek, W., Lozano, E., Mehlmann, G., Linnebank, F., Liem, J., ... André, E. (2012). Increasing Learners' Motivation through Pedagogical Agents: The Cast of Virtual Characters in the DynaLearn ILE. In *International workshop on agents for educational games and simulations '12* (pp. 151–165). (Cited on pp. 107, 108).
- Wolfe, G. (1994). *Shadow & Claw*. New York, NY: Orb Books. (Cited on p. xi).
- Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. (Cited on p. 127).
- Wouters, P., Oostendorp, H. V., Boonekamp, R. & van der Spek, E. D. (2011). The role of game discourse analysis and curiosity in creating engaging and effective serious games by implementing a back story and foreshadowing. *Interacting with Computers*, 23(4), 329–336. (Cited on p. 115).
- Wouters, P. & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, 60(1), 412–425. (Cited on p. 106).
- Wu, W.-H., Hsiao, H.-C., Wu, P.-L., Lin, C.-H. & Huang, S.-H. (2012). Investigating the learning-theory foundations of game-based learning: a meta-analysis. *Journal of Computer Assisted Learning*, 28(3), 265–279. (Cited on p. 10).
- Yannakakis, G. & Togelius, J. (2014). A Panorama of Artificial and Computational Intelligence in Games. *IEEE Transactions on Computational Intelligence and AI in Games*, 7(4), 317–335. (Cited on pp. 10, 167).
- Young, R. M., Riedl, M., Branly, M., Jhala, A., Martin, R. J. & Saretto, C. J. (2004). An architecture for integrating plan-based behavior generation with interactive game environments. *Journal of Game Development*, 1(1), 1–29. (Cited on pp. 19, 20).
- Zancanaro, M., Pianesi, F., Stock, O., Venuti, P., Cappelletti, A., Iandolo, G., ... Rossi, F. (2007). Children in the Museum: an Environment for Collaborative Storytelling. In *PEACH - Intelligent Interfaces for Museum Visits* (pp. 165–184). (Cited on p. 39).

PART V
APPENDICES

INTERACTIONS WITH THE INTERACTIVE STORYTELLER

A

Transcript 19 shows the original Dutch version of one of the stories generated by the children interacting with the Interactive Storyteller in the experiment by Alofs et al. (2015).

Er was eens een klein meisje met een rood kapje op haar hoofd. Ze wilde een verjaardagstaart naar haar oma brengen...

Wolfje is gemeen.

Roodkapje barst in lachen uit.

Oma bakt de kwarktaart.

Wolfje sluipt naar het huis van Roodkapje.

‘Hallo!’ zegt Roodkapje tegen Wolfje.

Met enkele korreltjes cyaankali vergiftigt Oma de kwarktaart.

‘Jij ook hallo...’ zegt Wolfje tegen Roodkapje.

Roodkapje doet een deel van de afwas.

Oma schuifelt naar het strand.

‘Geef hier!’ zegt Wolfje en pakt de verjaardagstaart van Roodkapje af.

‘Geef terug, die is van mij!’ zegt Roodkapje en pakt de verjaardagstaart weer terug van Wolfje.

Oma schuifelt naar het huis van Roodkapje.

‘Geef terug!’ zegt Wolfje en pakt de verjaardagstaart weer terug van Roodkapje.

‘Oh Oma luister!’ zegt Roodkapje. ‘Daarnet heeft Wolfje de verjaardagstaart van me gestolen!’

‘Och, och, wat vervelend,’ zegt Oma tegen Roodkapje.

Wolfje eet de verjaardagstaart op.

‘Alsjeblieft, voor jou!’ zegt Oma en geeft de kwarktaart aan Wolfje.

Roodkapje barst in lachen uit.

Oma barst in snikken uit.

Wolfje krijgt behoorlijk honger.

‘Wat moet ik doen?’ vraagt Roodkapje aan Oma.

‘Ok, ik heb een plan om Wolfje te gaan vergiftigen. Wacht maar af...’ zegt Oma tegen Roodkapje.

Wolfje eet de kwarktaart op.

Transcript 19.: The original Dutch version of a story created by two children with the Interactive Storyteller. See Transcript 3 for the English translation.

ADDITIONAL MATERIAL ON LOITER

B

This appendix contains additional material on LOITER-TB, LOITER-FB, and LOITER-AP.

B.1 LOITER-TB

Table 20 shows the utterances of which participants of the experiment on *loiter-tb* were asked to determine the stance (see Section 6.4). Table 21 shows the second SJT in that experiment, for which participants were asked to indicate which stance of the juvenile they would expect as a response on the shown utterances of a police officer. The subjective questionnaire in which participants rated to which extent they agreed with statements about their behaviour in the game is shown in Table 22.

Table 20.: The SJT on stance recognition for LOITER-TB.

UTTERANCE OF THE JUVENILE	INTENDED STANCE
'Ja, maar als wij dat doen, dan staat daar wel wat tegenover.'	DF
'OK OK, we doen wat je zegt.'	SF
'Ik kan hier niks mee, maar whatever.'	SH
'Ja doe, doe het lekker zelf joh!'	DH
'Ja, dat is heel duidelijk, dat snap ik wel.'	DF
'Als je het zo zegt... We zoeken wel een andere plek.'	SF
'Je snapt best wat ik bedoel...'	SH
'Jij komt hier echt alleen maar om te dissen, of niet?'	DH

Table 21.: The SJT on response estimation for LOITER-TB.

UTTERANCE OF THE POLICE OFFICER	RESPONDING STANCE
'Dit kunnen we samen makkelijk oplossen!'	SF
'Wat vind jij dan dat we moeten doen?'	DF
'Ik doe dit ook niet voor mn lol.'	DH
'Luister eens, we gaan doen wat ik zeg!'	SH
'Ja, dat is heel duidelijk, dat snap ik wel.'	SF
'Er is vast een andere oplossing voor te vinden, denk je niet?'	DF
'Dat wist ik niet, maar dat boeit me ook niet.'	DH
'Zo, jij gaat mooi op de bon, mannetje.'	SH

Table 22.: The subjective questionnaire for the experiment on LOITER-TB.

STATEMENT	RATING SCALE				
Ik wisselde vaak van aanpak.	1	2	3	4	5
Ik herkende de momenten waarop de jongere van houding veranderde.	1	2	3	4	5
Het scenario verliep zoals ik zou verwachten.	1	2	3	4	5
Het scenario verliep zoals ik het wilde.	1	2	3	4	5
Ik ben tevreden over de manier waarop ik het conflict heb opgelost.	1	2	3	4	5
Ik was in staat de houdingen van de jongere te onderscheiden.	1	2	3	4	5
Ik was in staat de verschillende keuzes die ik had te onderscheiden.	1	2	3	4	5
Ik snap nu, na het spelen van de game, beter hoe mijn gedrag dat van anderen beïnvloedt dan daarvoor.	1	2	3	4	5

Table 23.: The first part of the first situational judgement test (SJT1) for the experiment on LOITER-FB.

INTENDED STANCE	ACTION OF THE POLICE OFFICER	RATING SCALE						
DF	Je stapt rustig op de jongeren af. Je vraagt aan hen waarom ze mensen zo lastig vallen en drukt ze op het hart dat die vrouw dat echt niet fijn vond.	-3	-2	-1	0	1	2	3
SF	Je wacht even een minuut en stapt daarna rustig op de jongeren af. Je vraagt aan hen of zij wel eens uitgescholden zijn en hoe dat bij hen overkwam.	-3	-2	-1	0	1	2	3
SH	Je fluit naar de groep jongeren en schudt je hoofd. Daarna loop je langs hen op en gebaart dat ze maar beter ergens anders heen kunnen gaan.	-3	-2	-1	0	1	2	3
DH	Je stapt direct op de jongeren af en zegt dat ze zich normaal moeten gedragen. Dat mensen uitgescholden worden tolereer je niet. Anders is het kiezen: wegwezen of een boete.	-3	-2	-1	0	1	2	3

B.2 LOITER-FB

Tables 23 and 24 show the first situational judgement test (SJT1) used in the experiment on LOITER-FB (see Section 6.6). Participants were instructed to first read Transcripts 20 and 21, respectively, to introduce them to the setting. Table 25 shows SJT2; Transcript 22 shows the introductory text for this setting.

Een groepje hangjongeren hangt tegen een bankje in een winkelstraat. Jij staat iets verderop in de straat. Je ziet dat de jongeren staren naar een vrouw die voorbij komt gelopen. Als ze hen gepasseerd is, fluit één van de jongeren haar na. De vrouw kijkt met verbaasde blik om en de jongen zegt: ‘Hee, heb je zin om even mee te komen ofzo?’ De vrouw loopt snel door en negeert de groep, waarna de jongen haar na roept: ‘Slet!’ De hele groep barst hierop in lachen uit.

Transcript 20.: The first introductory text for SJT1.

Je houdt één van de jongeren die je laatst samen met zn vrienden hebt gesproken in de winkelstraat staande, omdat hij te hard reed op zijn scooter. Hij weet duidelijk dat hij erbij is en je ziet ook dat zijn achterlicht het niet goed doet. Hij kijkt je in eerste instantie niet aan, maar blikt weg zodra je op hem afstapt.

Transcript 21.: The second introductory text for SJT1.

Een groepje hangjongeren hangt tegen een bankje in een winkelstraat. Je loopt hen voorbij zonder er veel aandacht aan te schenken en ze laten ook niet merken dat ze je jou gezien hebben. Op het moment dat je langs hen bent gelopen roept één van hen met vervormde stem: ‘Kutwout!’ Daarna klinkt er wild gelach van de hele groep.

Transcript 22.: The introductory text for SJT2.

B.3 LOITER-AP

This section contains additional material on LOITER-AP.

B.3.1 *Pseudocode for profile stance calculation*

Listing 1 shows pseudocode of the algorithm governing the calculation of the dominance of a new stance.

Listing 1: Pseudocode for the calculation of the dominance of a new stance, given a profile and a player's stance.

```

// The p value determines to what extent the juvenile's dominance changes
p = 1.5

/* Calculates the dominance of a new stance d, given a stance profile
   ps and a player's stance αi */
calculate_dominance( ps, αi )
{
    // If the player is dominant (DF or DH)
    if( αi == DF || αi == DH )
    {
        // If the juvenile is dominant
        if( ps == DF || ps == DH )
            // The juvenile's dominance decreases with less than 1
            d = d -  $\frac{1}{p}$ 
        // If the juvenile is submissive
        else if( ps == SF || ps == SH )
            // The juvenile's dominance decreases with p
            d = d - p
    }
    // If the player is submissive (SF or SH)
    else if( αi == SF || αi == SH )
    {
        // If the juvenile is dominant
        if( ps == DF || ps == DH )
            // The juvenile's dominance increases with p
            d = d + p
        // If the juvenile is submissive
        else if( ps == SF || ps == SH )
            // The juvenile's dominance increases with less than 1
            d = d +  $\frac{1}{p}$ 
    }
}

```

B.3.2 Dutch translations of justifications

The original, Dutch justifications which were shown in the game are listed here. Table 26 lists the justifications shown in the thought bubbles; Table 27 lists those shown in the flashbacks.

B.3.3 *Explanation of game mechanics in LOITER-AP*

Figure 46 shows a screenshot of the explanation of the game mechanics of thought bubbles and flashbacks in LOITER-AP.

Figure 46.: The explanation of the thought bubbles and flashbacks in LOITER-AP.

Table 24.: The second part of the first situational judgement test (SJT1) for the experiment on LOITER-FB.

INTENDED STANCE	ACTION OF THE POLICE OFFICER	RATING SCALE						
DF	Je zegt tegen hem dat je denkt dat hij ook wel snapt wat er mis is. Je voegt daaraan toe dat je ziet dat zijn achterlicht kapot is. Je geeft hem echter alleen een bekeuring voor te hard rijden, zolang hij belooft het voortaan kalmer aan te doen en het achterlicht meteen te laten repareren.	-3	-2	-1	0	1	2	3
SF	Je vraagt of hij snapt wat er aan de hand is en legt uit dat een ongeluk in een klein hoekje zit. Ook maak je hem ervan bewust dat zijn achterlicht het niet doet en vraagt of hij dat snel laat repareren. Je geeft hem deze keer alleen een waarschuwing, als hij het maar rustig aan doet.	-3	-2	-1	0	1	2	3
SH	Je schudt je hoofd en kijkt eens naar het achterlicht. Je zegt dat dit er zo niet best uitziet en schrijft alvast een bon uit voor het te hard rijden. Je zegt dat je de moeite niet wilt nemen voor de tweede boete en stuurt de jongere weer op pad.	-3	-2	-1	0	1	2	3
DH	Je kijkt de jongere direct in de ogen aan en schrijft meteen twee bonnen uit. Je zegt dat hij maar beter op moeten passen in het verkeer, want met dit soort gedrag komen er zeker ongelukken van.	-3	-2	-1	0	1	2	3

Table 25.: The second situational judgement test (SJT2) for the experiment on LOITER-FB.

INTENDED STANCE	ACTION OF THE POLICE OFFICER	RATING SCALE						
DF	Je stapt rustig op de jongeren af. Je vraagt aan hen waarom ze mensen zo lastig vallen en drukt ze op het hart dat die vrouw dat echt niet fijn vond.	-3	-2	-1	0	1	2	3
SF	Je wacht even een minuut en stapt daarna rustig op de jongeren af. Je vraagt aan hen of zij wel eens uitgescholden zijn en hoe dat bij hen overkwam.	-3	-2	-1	0	1	2	3
SH	Je fluit naar de groep jongeren en schudt je hoofd. Daarna loop je langs hen op en gebaart dat ze maar beter ergens anders heen kunnen gaan.	-3	-2	-1	0	1	2	3
DH	Je stapt direct op de jongeren af en zegt dat ze zich normaal moeten gedragen. Dat mensen uitgescholden worden tolereer je niet. Anders is het kiezen: wegwezen of een boete.	-3	-2	-1	0	1	2	3

Table 26.: The original, Dutch justifications that can be shown in thought bubbles when a within-scenario adaptation is made. See Table 14 for the English translation.

SCENARIO	INTENDED STANCE	JUSTIFICATION
1	DH	‘Het mannetje zal eens zeggen wat we moeten doen zeker. Oprotten met die wout!’
	SH	‘Damn, wat een wuss. Geen aandacht besteden aan die loser.’
2	DH	‘Wat zoekt die vent ruzie? Hij moet echt niet met me fokken man!’
	SH	‘Meen je toch niet, mister nice guy komt ons lastig vallen. Fok deze shit.’
3	DH	‘Wat een commandant, we zitten niet in het leger! Zoekt-ie ruzie ofzo, kom maar op!’
	SH	‘Shit, hij heeft er zin in... Fuck, liet-ie ons maar met rust.’

Table 27.: The original, Dutch justifications that can be shown in flashbacks when a between-scenario adaptation is made. See Table 15 for the English translation.

SCENARIO	INTENDED STANCE	JUSTIFICATION
2	DH	‘Vanochtend deed ik fucking alles fout in de supermarkt, nu komt deze vent ook nog zeiken? Ze moeten me gewoon eens met rust laten!’
	SH	‘Fuck, vanochtend viel m’n manager me lastig tijdens het vakken vullen, nu oom agent weer. Wat een fucking gedoe...’
3	DH	‘Telefoon kwijt aan m’n ouders, nu komt deze vent me hier lastig vallen?! Ze moeten kappen met mij fokken!’
	SH	‘Oh shit, ’t wordt steeds beter. Eerst m’n telefoon ingepikt door m’n ouders, nu dit. Fuck, heb hier echt geen zin in...’

SIKS DISSERTATIEREEKS

2011

2011-01 Botond Cseke (RUN) *Variational Algorithms for Bayesian Inference in Latent Gaussian Models*

2011-02 Nick Tinnemeier(UU) *Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language*

2011-03 Jan Martijn van der Werf (TUE) *Compositional Design and Verification of Component-Based Information Systems*

2011-04 Hado van Hasselt (UU) *Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference learning algorithms*

2011-05 Base van der Raadt (VU) *Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.*

2011-06 Yiwen Wang (TUE) *Semantically-Enhanced Recommendations in Cultural Heritage*

2011-07 Yujia Cao (UT) *Multimodal Information Presentation for High Load Human Computer Interaction*

2011-08 Nieske Vergunst (UU) *BDI-based Generation of Robust Task-Oriented Dialogues*

2011-09 Tim de Jong (OU) *Contextualised Mobile Media for Learning*

2011-10 Bart Bogaert (UvT) *Cloud Content Contention*

2011-11 Dhaval Vyas (UT) *Designing for Awareness: An Experience-focused HCI Perspective*

2011-12 Carmen Bratosin (TUE) *Grid Architecture for Distributed Process Mining*

2011-13 Xiaoyu Mao (UvT) *Airport under Control. Multiagent Scheduling for Airport Ground Handling*

2011-14 Milan Lovric (EUR) *Behavioral Finance and Agent-Based Artificial Markets*

2011-15 Marijn Koolen (UvA) *The Meaning of Structure: the Value of Link Evidence for Information Retrieval*

2011-16 Maarten Schadd (UM) *Selective Search in Games of Different Complexity*

2011-17 Jiyin He (UvA) *Exploring Topic Structure: Coherence, Diversity and Relatedness*

2011-18 Mark Ponsen (UM) *Strategic Decision-Making in complex games*

2011-19 Ellen Rusman (OU) *The Mind's Eye on Personal Profiles*

2011-20 Qing Gu (VU) *Guiding service-oriented software engineering - A view-based approach*

2011-21 Linda Terlouw (TUD) *Modularization and Specification of Service-Oriented Systems*

2011-22 Junte Zhang (UvA) *System Evaluation of Archival Description and Access*

2011-23 Wouter Weerkamp (UvA) *Finding People and their Utterances in Social Media*

2011-24 Herwin van Welbergen (UT) *Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior*

2011-25 Syed Waqar ul Qounain Jaffry (VU) *Analysis and Validation of Models for Trust Dynamics*

2011-26 Matthijs Aart Pontier (VU) *Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots*

2011-27 Aniel Bhulai (VU) *Dynamic website optimization through autonomous management of design patterns*

2011-28 Rianne Kaptein(UvA) *Effective Focused Retrieval by Exploiting Query Context and Document Structure*

2011-29 Faisal Kamiran (TUE) *Discrimination-aware Classification*

2011-30 Egon van den Broek (UT) *Affective Signal Processing (ASP): Unraveling the mystery of emotions*

2011-31 Ludo Waltman (EUR) *Computational and Game-Theoretic Approaches for Modeling Bounded Rationality*

2011-32 Nees-Jan van Eck (EUR) *Methodological Advances in Bibliometric Mapping of Science*

2011-33 Tom van der Weide (UU) *Arguing to Motivate Decisions*

2011-34 Paolo Turrini (UU) *Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations*

2011-35 Maaïke Harbers (UU) *Explaining Agent Behavior in Virtual Training*

2011-36 Erik van der Spek (UU) *Experiments in serious game design: a cognitive approach*

2011-37 Adriana Burlutiu (RUN) *Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference*

2011-38 Nyree Lemmens (UM) *Bee-inspired Distributed Optimization*

2011-39 Joost Westra (UU) *Organizing Adaptation using Agents in Serious Games*

2011-40 Viktor Clerc (VU) *Architectural Knowledge Management in Global Software Development*

2011-41 Luan Ibraimi (UT) *Cryptographically Enforced Distributed Data Access Control*

2011-42 Michal Sindlar (UU) *Explaining Behavior through Mental State Attribution*

2011-43 Henk van der Schuur (UU) *Process Improvement through Software Operation Knowledge*

2011-44 Boris Reuderink (UT) *Robust Brain-Computer Interfaces*

2011-45 Herman Stehouwer (UvT) *Statistical Language Models for Alternative Sequence Selection*

2011-46 Beibei Hu (TUD) *Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work*

2011-47 Azizi Bin Ab Aziz(VU) *Exploring Computational Models for Intelligent Support of Persons with Depression*

2011-48 Mark Ter Maat (UT) *Response Selection and Turn-taking for a Sensitive Artificial Listening Agent*

2011-49 Andreea Niculescu (UT) *Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality*

2012

2012-01 Terry Kakeeto (UvT) *Relationship Marketing for SMEs in Uganda*

2012-02 Muhammad Umair(VU) *Adaptivity, emotion, and Rationality in Human and Ambient Agent Models*

2012-03 Adam Vanya (VU) *Supporting Architecture Evolution by Mining Software Repositories*

2012-04 Jurriaan Souer (UU) *Development of Content Management System-based Web Applications*

2012-05 Marijn Plomp (UU) *Maturing Interorganisational Information Systems*

2012-06 Wolfgang Reinhardt (OU) *Awareness Support for Knowledge Workers in Research Networks*

- 2012-07** Rianne van Lambalgen (VU) *When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions*
- 2012-08** Gerben de Vries (UVA) *Kernel Methods for Vessel Trajectories*
- 2012-09** Ricardo Neisse (UT) *Trust and Privacy Management Support for Context-Aware Service Platforms*
- 2012-10** David Smits (TUE) *Towards a Generic Distributed Adaptive Hypermedia Environment*
- 2012-11** J.C.B. Rantham Prabhakara (TUE) *Process Mining in the Large: Preprocessing, Discovery, and Diagnostics*
- 2012-12** Kees van der Sluijs (TUE) *Model Driven Design and Data Integration in Semantic Web Information Systems*
- 2012-13** Suleman Shahid (UvT) *Fun and Face: Exploring non-verbal expressions of emotion during playful interactions*
- 2012-14** Evgeny Knutov(TUE) *Generic Adaptation Framework for Unifying Adaptive Web-based Systems*
- 2012-15** Natalie van der Wal (VU) *Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes.*
- 2012-16** Fiemke Both (VU) *Helping people by understanding them - Ambient Agents supporting task execution and depression treatment*
- 2012-17** Amal Elgammal (UvT) *Towards a Comprehensive Framework for Business Process Compliance*
- 2012-18** Eltjo Poort (VU) *Improving Solution Architecting Practices*
- 2012-19** Helen Schonenberg (TUE) *What's Next? Operational Support for Business Process Execution*
- 2012-20** Ali Bahramisharif (RUN) *Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing*
- 2012-21** Roberto Cornacchia (TUD) *Querying Sparse Matrices for Information Retrieval*
- 2012-22** Thijs Vis (UvT) *Intelligence, politie en veiligheidsdienst: verenigbare grootheden?*
- 2012-23** Christian Muehl (UT) *Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction*
- 2012-24** Laurens van der Werff (UT) *Evaluation of Noisy Transcripts for Spoken Document Retrieval*
- 2012-25** Silja Eckartz (UT) *Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application*
- 2012-26** Emile de Maat (UVA) *Making Sense of Legal Text*
- 2012-27** Hayrettin Gurkok (UT) *Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games*
- 2012-28** Nancy Pascall (UvT) *Engendering Technology Empowering Women*
- 2012-29** Almer Tigelaar (UT) *Peer-to-Peer Information Retrieval*
- 2012-30** Alina Pommeranz (TUD) *Designing Human-Centered Systems for Reflective Decision Making*
- 2012-31** Emily Bagarukayo (RUN) *A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure*
- 2012-32** Wietske Visser (TUD) *Qualitative multi-criteria preference representation and reasoning*
- 2012-33** Rory Sie (OUN) *Coalitions in Cooperation Networks (CO-COON)*
- 2012-34** Pavol Jancura (RUN) *Evolutionary analysis in PPI networks and applications*
- 2012-35** Evert Haasdijk (VU) *Never Too Old To Learn - On-line Evolution of Controllers in Swarm- and Modular Robotics*
- 2012-36** Denis Ssebugwawo (RUN) *Analysis and Evaluation of Collaborative Modeling Processes*
- 2012-37** Agnes Nakakawa (RUN) *A Collaboration Process for Enterprise Architecture Creation*
- 2012-38** Selmar Smit (VU) *Parameter Tuning and Scientific Testing in Evolutionary Algorithms*
- 2012-39** Hassan Fatemi (UT) *Risk-aware design of value and co-ordination networks*
- 2012-40** Agus Gunawan (UvT) *Information Access for SMEs in Indonesia*
- 2012-41** Sebastian Kelle (OU) *Game Design Patterns for Learning*
- 2012-42** Dominique Verpoorten (OU) *Reflection Amplifiers in self-regulated Learning*
- 2012-43** Withdrawn
- 2012-44** Anna Tordai (VU) *On Combining Alignment Techniques*
- 2012-45** Benedikt Kratz (UvT) *A Model and Language for Business-aware Transactions*
- 2012-46** Simon Carter (UVA) *Exploration and Exploitation of Multilingual Data for Statistical Machine Translation*
- 2012-47** Manos Tsagkias (UVA) *Mining Social Media: Tracking Content and Predicting Behavior*
- 2012-48** Jorn Bakker (TUE) *Handling Abrupt Changes in Evolving Time-series Data*
- 2012-49** Michael Kaisers (UM) *Learning against Learning - Evolutionary dynamics of reinforcement learning algorithms in strategic interactions*
- 2012-50** Steven van Kervel (TUD) *Ontology driven Enterprise Information Systems Engineering*
- 2012-51** Jeroen de Jong (TUD) *Heuristics in Dynamic Sceduling; a practical framework with a case study in elevator dispatching*
- 2013**
- 2013-01** Viorel Milea (EUR) *News Analytics for Financial Decision Support*
- 2013-02** Erietta Liarou (CWI) *MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing*
- 2013-03** Szymon Klarman (VU) *Reasoning with Contexts in Description Logics*
- 2013-04** Chetan Yadati(TUD) *Coordinating autonomous planning and scheduling*
- 2013-05** Dulce Pumareja (UT) *Groupware Requirements Evolutions Patterns*
- 2013-06** Romulo Goncalves(CWI) *The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience*
- 2013-07** Giel van Lankveld (UvT) *Quantifying Individual Player Differences*
- 2013-08** Robbert-Jan Merk(VU) *Making enemies: cognitive modeling for opponent agents in fighter pilot simulators*
- 2013-09** Fabio Gori (RUN) *Metagenomic Data Analysis: Computational Methods and Applications*
- 2013-10** Jeewanie Jayasinghe Arachchige(UvT) *A Unified Modeling Framework for Service Design.*
- 2013-11** Evangelos Pournaras(TUD) *Multi-level Reconfigurable Self-organization in Overlay Services*
- 2013-12** Marian Razavian(VU) *Knowledge-driven Migration to Services*
- 2013-13** Mohammad Safiri(UT) *Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly*
- 2013-14** Jafar Tanha (UVA) *Ensemble Approaches to Semi-Supervised Learning Learning*
- 2013-15** Daniel Hennes (UM) *Multiagent Learning - Dynamic Games and Applications*
- 2013-16** Eric Kok (UU) *Exploring the practical benefits of argumentation in multi-agent deliberation*
- 2013-17** Koen Kok (VU) *The PowerMatcher: Smart Coordination for the Smart Electricity Grid*

- 2013-18** Jeroen Janssens (UvT) *Outlier Selection and One-Class Classification*
- 2013-19** Renze Steenhuizen (TUD) *Coordinated Multi-Agent Planning and Scheduling*
- 2013-20** Katja Hofmann (UvA) *Fast and Reliable Online Learning to Rank for Information Retrieval*
- 2013-21** Sander Wubben (UvT) *Text-to-text generation by monolingual machine translation*
- 2013-22** Tom Claassen (RUN) *Causal Discovery and Logic*
- 2013-23** Patricio de Alencar Silva (UvT) *Value Activity Monitoring*
- 2013-24** Haitham Bou Ammar (UM) *Automated Transfer in Reinforcement Learning*
- 2013-25** Agnieszka Anna Latoszek-Berendsen (UM) *Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System*
- 2013-26** Alireza Zarghami (UT) *Architectural Support for Dynamic Homecare Service Provisioning*
- 2013-27** Mohammad Huq (UT) *Inference-based Framework Managing Data Provenance*
- 2013-28** Frans van der Sluis (UT) *When Complexity becomes Interesting: An Inquiry into the Information eXperience*
- 2013-29** Iwan de Kok (UT) *Listening Heads*
- 2013-30** Joyce Nakatumba (TUE) *Resource-Aware Business Process Management: Analysis and Support*
- 2013-31** Dinh Khoa Nguyen (UvT) *Blueprint Model and Language for Engineering Cloud Applications*
- 2013-32** Kamakshi Rajagopal (OUN) *Networking For Learning; The role of Networking in a Lifelong Learner's Professional Development*
- 2013-33** Qi Gao (TUD) *User Modeling and Personalization in the Microblogging Sphere*
- 2013-34** Kien Tjin-Kam-Jet (UT) *Distributed Deep Web Search*
- 2013-35** Abdallah El Ali (UvA) *Minimal Mobile Human Computer Interaction Promotor: Prof. dr. L. Hardman (CWI/UvA)*
- 2013-36** Than Lam Hoang (TUE) *Pattern Mining in Data Streams*
- 2013-37** Dirk Börner (OUN) *Ambient Learning Displays*
- 2013-38** Eelco den Heijer (VU) *Autonomous Evolutionary Art*
- 2013-39** Joop de Jong (TUD) *A Method for Enterprise Ontology based Design of Enterprise Information Systems*
- 2013-40** Pim Nijssen (UM) *Monte-Carlo Tree Search for Multi-Player Games*
- 2013-41** Jochem Liem (UvA) *Supporting the Conceptual Modeling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning*
- 2013-42** Léon Planken (TUD) *Algorithms for Simple Temporal Reasoning*
- 2013-43** Marc Bron (UvA) *Exploration and Contextualization through Interaction and Concepts*
- 2014**
- 2014-01** Nicola Barile (UU) *Studies in Learning Monotone Models from Data*
- 2014-02** Fiona Tuliayo (RUN) *Combining System Dynamics with a Domain Modeling Method*
- 2014-03** Sergio Raul Duarte Torres (UT) *Information Retrieval for Children: Search Behavior and Solutions*
- 2014-04** Hanna Jochmann-Mannak (UT) *Websites for children: search strategies and interface design - Three studies on children's search performance and evaluation*
- 2014-05** Jurriaan van Reijssen (UU) *Knowledge Perspectives on Advancing Dynamic Capability*
- 2014-06** Damian Tamburri (VU) *Supporting Networked Software Development*
- 2014-07** Arya Adriansyah (TUE) *Aligning Observed and Modeled Behavior*
- 2014-08** Samur Araujo (TUD) *Data Integration over Distributed and Heterogeneous Data Endpoints*
- 2014-09** Philip Jackson (UvT) *Toward Human-Level Artificial Intelligence: Representation and Computation of Meaning in Natural Language*
- 2014-10** Ivan Salvador Razo Zapata (VU) *Service Value Networks*
- 2014-11** Janneke van der Zwaan (TUD) *An Empathic Virtual Buddy for Social Support*
- 2014-12** Willem van Willigen (VU) *Look Ma, No Hands: Aspects of Autonomous Vehicle Control*
- 2014-13** Arlette van Wissen (VU) *Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains*
- 2014-14** Yangyang Shi (TUD) *Language Models With Meta-information*
- 2014-15** Natalya Mogles (VU) *Agent-Based Analysis and Support of Human Functioning in Complex Socio-Technical Systems: Applications in Safety and Healthcare*
- 2014-16** Krystyna Milian (VU) *Supporting trial recruitment and design by automatically interpreting eligibility criteria*
- 2014-17** Kathrin Dentler (VU) *Computing healthcare quality indicators automatically: Secondary Use of Patient Data and Semantic Interoperability*
- 2014-18** Mattijs Ghijsen (UvA) *Methods and Models for the Design and Study of Dynamic Agent Organizations*
- 2014-19** Vinicius Ramos (TUE) *Adaptive Hypermedia Courses: Qualitative and Quantitative Evaluation and Tool Support*
- 2014-20** Mena Habib (UT) *Named Entity Extraction and Disambiguation for Informal Text: The Missing Link*
- 2014-21** Cassidy Clark (TUD) *Negotiation and Monitoring in Open Environments*
- 2014-22** Marieke Peeters (UU) *Personalized Educational Games - Developing agent-supported scenario-based training*
- 2014-23** Eleftherios Sidiourgos (UvA/CWI) *Space Efficient Indexes for the Big Data Era*
- 2014-24** Davide Ceolin (VU) *Trusting Semi-structured Web Data*
- 2014-25** Martijn Lappenschaar (RUN) *New network models for the analysis of disease interaction*
- 2014-26** Tim Baarslag (TUD) *What to Bid and When to Stop*
- 2014-27** Rui Jorge Almeida (EUR) *Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty*
- 2014-28** Anna Chmielowiec (VU) *Decentralized k-Clique Matching*
- 2014-29** Jaap Kabbedijk (UU) *Variability in Multi-Tenant Enterprise Software*
- 2014-30** Peter de Cock (UvT) *Anticipating Criminal Behaviour*
- 2014-31** Leo van Moergestel (UU) *Agent Technology in Agile Multiparallel Manufacturing and Product Support*
- 2014-32** Naser Ayat (UvA) *On Entity Resolution in Probabilistic Data*
- 2014-33** Tesfa Tegegne (RUN) *Service Discovery in eHealth*
- 2014-34** Christina Manteli (VU) *The Effect of Governance in Global Software Development: Analyzing Transactive Memory Systems*
- 2014-35** Joost van Ooijen (UU) *Cognitive Agents in Virtual Worlds: A Middleware Design Approach*
- 2014-36** Joos Buijs (TUE) *Flexible Evolutionary Algorithms for Mining Structured Process Models*
- 2014-37** Maral Dadvar (UT) *Experts and Machines United Against Cyberbullying*
- 2014-38** Danny Plass-Oude Bos (UT) *Making brain-computer interfaces better: improving usability through post-processing*
- 2014-39** Jasmina Maric (UvT) *Web Communities, Immigration, and Social Capital*

- 2014-40** Walter Omona (RUN) *A Framework for Knowledge Management Using ICT in Higher Education*
- 2014-41** Frederic Hogenboom (EUR) *Automated Detection of Financial Events in News Text*
- 2014-42** Carsten Eijckhof (CWI/TUD) *Contextual Multidimensional Relevance Models*
- 2014-43** Kevin Vlaanderen (UU) *Supporting Process Improvement using Method Increments*
- 2014-44** Paulien Meesters (UvT) *Intelligent Blauw. Met als ondertitel: Intelligent-gestuurde politiezorg in gebiedsgebonden eenheden.*
- 2014-45** Birgit Schmitz (OUN) *Mobile Games for Learning: A Pattern-Based Approach*
- 2014-46** Ke Tao (TUD) *Social Web Data Analytics: Relevance, Redundancy, Diversity*
- 2014-47** Shangsong Liang (UVA) *Fusion and Diversification in Information Retrieval*
- 2015**
- 2015-01** Niels Netten (UvA) *Machine Learning for Relevance of Information in Crisis Response*
- 2015-02** Faiza Bukhsh (UvT) *Smart auditing: Innovative Compliance Checking in Customs Controls*
- 2015-03** Twan van Laarhoven (RUN) *Machine learning for network data*
- 2015-04** Howard Spoelstra (OUN) *Collaborations in Open Learning Environments*
- 2015-05** Christoph Bösch(UT) *Cryptographically Enforced Search Pattern Hiding*
- 2015-06** Farideh Heidari (TUD) *Business Process Quality Computation – Computing Non-Functional Requirements to Improve Business Processes*
- 2015-07** Maria-Hendrike Peetz(UvA) *Time-Aware Online Reputation Analysis*
- 2015-08** Jie Jiang (TUD) *Organizational Compliance: An agent-based model for designing and evaluating organizational interactions*
- 2015-09** Randy Klaassen(UT) *HCI Perspectives on Behavior Change Support Systems*
- 2015-10** Henry Hermans (OUN) *OpenU: design of an integrated system to support lifelong learning*
- 2015-11** Yongming Luo(TUE) *Designing algorithms for big graph datasets: A study of computing bisimulation and joins*
- 2015-12** Julie M. Birkholz (VU) *Modi Operandi of Social Network Dynamics: The Effect of Context on Scientific Collaboration Networks*
- 2015-13** Giuseppe Procaccianti(VU) *Energy-Efficient Software*
- 2015-14** Bart van Straalen (UT) *A cognitive approach to modeling bad news conversations*
- 2015-15** Klaas Andries de Graaf (VU) *Ontology-based Software Architecture Documentation*
- 2015-16** Changyun Wei (UT) *Cognitive Coordination for Cooperative Multi-Robot Teamwork*
- 2015-17** André van Cleeff (UT) *Physical and Digital Security Mechanisms: Properties, Combinations and Trade-offs*
- 2015-18** Holger Pirk (CWI) *Waste Not, Want Not! Managing Relational Data in Asymmetric Memories*
- 2015-19** Bernardo Tabuenca (OUN) *Ubiquitous Technology for Lifelong Learners*
- 2015-20** Lois Vanhée (UU) *Using Culture and Values to Support Flexible Coordination*
- 2015-21** Sibren Fetter (OUN) *Using Peer-Support to Expand and Stabilize Online Learning*
- 2015-22** Zhemine Zhu(UT) *Co-occurrence Rate Networks*
- 2015-23** Luit Gazendam (VU) *Cataloguer Support in Cultural Heritage*
- 2015-24** Richard Berendsen (UVA) *Finding People, Papers, and Posts: Vertical Search Algorithms and Evaluation*
- 2015-25** Steven Woudenberg (UU) *Bayesian Tools for Early Disease Detection*
- 2015-26** Alexander Hogenboom (EUR) *Sentiment Analysis of Text Guided by Semantics and Structure*
- 2015-27** Sándor Héman (CWI) *Updating compressed column stores*
- 2015-28** Janet Bagorogozo (TIU) *KNOWLEDGE MANAGEMENT AND HIGH PERFORMANCE; The Uganda Financial Institutions Model for HPO*
- 2015-29** Hendrik Baier (UM) *Monte-Carlo Tree Search Enhancements for One-Player and Two-Player Domains*
- 2015-30** Kiavash Bahreini (OU) *Real-time Multimodal Emotion Recognition in E-Learning*
- 2015-31** Yakup Koç (TUD) *On the robustness of Power Grids*
- 2015-32** Jerome Gard (UL) *Corporate Venture Management in SMEs*
- 2015-33** Frederik Schadd (TUD) *Ontology Mapping with Auxiliary Resources*
- 2015-34** Victor de Graaf(UT) *Gesocial Recommender Systems*
- 2015-35** Jungxao Xu (TUD) *Affective Body Language of Humanoid Robots: Perception and Effects in Human Robot Interaction*
- 2016**
- 2016-01** Syed Saiden Abbas (RUN) *Recognition of Shapes by Humans and Machines*
- 2016-02** Michiel Christiaan Meulendijk (UU) *Optimizing medication reviews through decision support: prescribing a better pill to swallow*
- 2016-03** Maya Sappelli (RUN) *Knowledge Work in Context: User Centered Knowledge Worker Support*
- 2016-04** Laurens Rietveld (VU) *Publishing and Consuming Linked Data*
- 2016-05** Evgeny Sherkhonov (UVA) *Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers*
- 2016-06** Michel Wilson (TUD) *Robust scheduling in an uncertain environment*
- 2016-07** Jeroen de Man (VU) *Measuring and modeling negative emotions for virtual training*
- 2016-08** Matje van de Camp (TiU) *A Link to the Past: Constructing Historical Social Networks from Unstructured Data*
- 2016-09** Archana Nottamkandath (VU) *Trusting Crowdsourced Information on Cultural Artefacts*
- 2016-10** George Karafotias (VUA) *Parameter Control for Evolutionary Algorithms*
- 2016-11** Anne Schuth (UVA) *Search Engines that Learn from Their Users*
- 2016-12** Max Knobbout (UU) *Logics for Modelling and Verifying Normative Multi-Agent Systems*
- 2016-13** Nana Baah Gyan (VU) *The Web, Speech Technologies and Rural Development in West Africa – An ICT4D Approach*
- 2016-14** Ravi Khadka (UU) *Revisiting Legacy Software System Modernization*
- 2016-15** Steffen Michels (RUN) *Hybrid Probabilistic Logics – Theoretical Aspects, Algorithms and Experiments*
- 2016-16** Guangliang Li (UVA) *Socially Intelligent Autonomous Agents that Learn from Human Reward*
- 2016-17** Berend Weel (VU) *Towards Embodied Evolution of Robot Organisms*
- 2016-18** Albert Meroño Peñuela (VU) *Refining Statistical Data on the Web*
- 2016-19** Julia Efremova (Tu/e) *Mining Social Structures from Genealogical Data*
- 2016-20** Daan Odijk (UVA) *Context & Semantics in News & Web Search*

- 2016-21** Alejandro Moreno Céleri (UT) *From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground*
- 2016-22** Grace Lewis (VU) *Software Architecture Strategies for Cyber-Foraging Systems*
- 2016-23** Fei Cai (UVA) *Query Auto Completion in Information Retrieval*
- 2016-24** Brend Wanders (UT) *Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach*
- 2016-25** Julia Kiseleva (TU/e) *Using Contextual Information to Understand Searching and Browsing Behavior*
- 2016-26** Dilhan Thilakarathne (VU) *In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains*
- 2016-27** Wen Li (TUD) *Understanding Geo-spatial Information on Social Media*
- 2016-28** Mingxin Zhang (TUD) *Large-scale Agent-based Social Simulation - A study on epidemic prediction and control*
- 2016-29** Nicolas Höning (TUD) *Peak reduction in decentralised electricity systems - Markets and prices for flexible planning*
- 2016-30** Ruud Mattheij (UvT) *The Eyes Have It*
- 2016-31** Mohammad Khelghati (UT) *Deep web content monitoring*
- 2016-32** Eelco Vriezেকolk (UT) *Assessing Telecommunication Service Availability Risks for Crisis Organisations*
- 2016-33** Peter Bloem (UVA) *Single Sample Statistics, exercises in learning from just one example*
- 2016-34** Dennis Schunselaar (TU/e) *Configurable Process Trees: Elicitation, Analysis, and Enactment*
- 2016-35** Zhaochun Ren (UVA) *Monitoring Social Media: Summarization, Classification and Recommendation*
- 2016-36** Daphne Karreman (UT) *Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies*
- 2016-37** Giovanni Sileno (UvA) *Aligning Law and Action - a conceptual and computational inquiry*
- 2016-38** Andrea Minuto (UT) *Materials that Matter - Smart Materials meet Art & Interaction Design*
- 2016-39** Merijn Bruijnes (UT) *Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect*
- 2016-40** Christian Detweiler (TUD) *Accounting for Values in Design*
- 2016-41** Thomas King (TUD) *Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance*
- 2016-42** Spyros Martzoukos (UVA) *Combinatorial and Compositional Aspects of Bilingual Aligned Corpora*
- 2016-43** Saskia Koldijk (RUN) *Context-Aware Support for Stress Self-Management: From Theory to Practice*
- 2016-44** Thibault Sellam (UVA) *Automatic Assistants for Database Exploration*
- 2016-45** Bram van de Laar (UT) *Experiencing Brain-Computer Interface Control*
- 2016-46** Jorge Gallego Perez (UT) *Robots to Make you Happy*
- 2016-47** Christina Weber (UL) *Real-time foresight - Preparedness for dynamic innovation networks*
- 2016-48** Tanja Buttler (TUD) *Collecting Lessons Learned*
- 2016-49** Gleb Polevoy (TUD) *Participation and Interaction in Projects. A Game-Theoretic Analysis*
- 2016-50** Yan Wang (UVT) *The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains*
- 2017**
- 2017-01** Jan-Jaap Oerlemans (UL) *Investigating Cybercrime*
- 2017-02** Sjoerd Timmer (UU) *Designing and Understanding Forensic Bayesian Networks using Argumentation*
- 2017-03** Daniël Harold Telgen (UU) *Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines*
- 2017-04** Mrunal Gawade (CWI) *Multi-core Parallelism in a Column-Store*
- 2017-05** Mahdiah Shadi (UVA) *Collaboration Behavior*
- 2017-06** Damir Vandić (EUR) *Intelligent Information Systems for Web Product Search*
- 2017-07** Roel Bertens (UU) *Insight in Information: from Abstract to Anomaly*
- 2017-08** Rob Konijn (VU) *Detecting Interesting Differences: Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery*
- 2017-09** Dong Nguyen (UT) *Text as Social and Cultural Data: A Computational Perspective on Variation in Text*
- 2017-10** Robby van Delden (UT) *(Steering) Interactive Play Behavior*
- 2017-11** Florian Kunneman (RUN) *Modelling patterns of time and emotion in Twitter #anticipointment*
- 2017-12** Sander Leemans (UT) *Robust Process Mining with Guarantees*
- 2017-13** Gijs Huisman (UT) *Social Touch Technology - Extending the reach of social touch through haptic technology*
- 2017-14** Shoshannah Tekofsky (UvT) *You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior*
- 2017-15** Peter Berck, Radboud University (RUN) *Memory-Based Text Correction*
- 2017-16** Aleksandr Chuklin (UVA) *Understanding and Modeling Users of Modern Search Engines*
- 2017-17** Daniel Dimov (VU) *Crowdsourced Online Dispute Resolution*
- 2017-18** Ridho Reinanda (UVA) *Entity Associations for Search*
- 2017-19** Jeroen Vuurens (TUD) *Proximity of Terms, Texts and Semantic Vectors in Information Retrieval*
- 2017-20** Mohammadbashir Sedighi (TUD) *Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility*
- 2017-21** Jeroen Linssen (UT) *Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)*

STELLINGEN

behorende bij het proefschrift:

Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)

in het openbaar te verdedigen door Johannes Maria Linssen

op donderdag 4 mei, 2017 om 16:45 uur

1. Verhalen, zowel klassiek als interactief, kunnen op verschillende manieren ervaren worden. Juist de momenten waarop lezers van perspectief veranderen, zijn de momenten die het meest over het verhaalsysteem en de ervaring van de lezer onthullen. *(Hoofdstuk 3)*
2. Drama ontspringt aan conflict. Een verhaalwereld zonder conflicten tussen karakters is een wereld zonder drama. Het doel van een verhaalgeneratiesysteem is derhalve het creëren van conflicten. *(Hoofdstuk 4)*
3. De moeilijkheid van sociale interactie zit hem in het herkennen van de redenen waardoor iemands gedrag verandert. In een educatief spel kan de juiste feedback deze redenen benadrukken en de speler inzicht in de interpersoonlijke verhoudingen van karakters verschaffen. *(Hoofdstuk 6)*
4. Ingrijpende, onverwachte wijzigingen in het gedrag van karakters, en daardoor in de loop van een verhaal, kunnen met de juiste woorden weer coherent gemaakt worden. *(Hoofdstuk 7)*
5. Van Gödel kunnen wij leren dat men voor waarheidsvinding nooit meta genoeg kan zijn.
6. Alle games worden serieus als ze als zodanig benaderd worden.
7. Als context alles is, en wetenschap zaken modelleert door context te verwijderen, zegt wetenschap niets.
8. Niets veroudert zo snel als de toekomst; in de wetenschap veroudert niets zo snel als een nieuw inzicht.
9. De hele wereld is een spel / En wij zijn allen niets meer dan gamers
10. Metal ben je niet voor even; metal ben je voor het leven.

PROPOSITIONS

accompanying the dissertation:

Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)

to be publicly defended by Johannes Maria Linssen

on Thursday May 4, 2017 at 16:45

1. Stories, both classical and interactive, can be experienced in different ways. The moments during which readers change perspective reveal the most about the storytelling system and the reader's experience. (*Chapter 3*)
2. Drama exists through conflict. A storyworld without conflicts between characters is a world without drama. Therefore, the goal of a story generation system is the creation of conflicts. (*Chapter 4*)
3. The difficulty of social interaction is found in recognizing the reasons causing changes in someone's behaviour. In a serious game, the right feedback can emphasize these reasons and grant the player insight into the interpersonal relations between characters. (*Chapter 6*)
4. Drastic, unexpected changes in the behaviour of characters, and hence in the course of a story, can be made coherent again with the right words. (*Chapter 7*)
5. From Gödel we can learn that, to find truth, one can never be meta enough.
6. All games become serious if they are approached as such.
7. If context is everything, and science models affairs by removing context, science says nothing.
8. Nothing ages as fast as the future; in science, nothing ages as fast as a new insight.
9. All the world's a game / And all the men and women merely gamers
10. Metal is not simply rife; metal is what you are for life.