Department of Computer Science
Series of Publications A
Report A-2019-3

Human-Computer Co-Creativity — Designing, Evaluating and Modelling Computational Collaborators for Poetry Writing

Anna Kantosalo

Doctoral dissertation, to be presented for public discussion with the permission of the Faculty of Science of the University of Helsinki, in Auditorium CK112, Exactum building, on the 13th of August, 2019 at 12 o'clock.

University of Helsinki
Finland

Supervisor
Hannu Toivonen, University of Helsinki, Finland

Pre-examiners
Georgios N. Yannakakis, University of Malta, Malta
Brian Magerko, Ivan Allen College, Georgia Institute of Technology, Georgia, United States of America

Opponent
Mary Lou Maher, University of North Carolina at Charlotte, North Carolina, United States of America

Custos
Hannu Toivonen, University of Helsinki, Finland

Contact information

Department of Computer Science
P.O. Box 68 (Pietari Kalmin katu 5)
FI-00014 University of Helsinki
Finland

Email address: info@cs.helsinki.fi
URL: http://cs.helsinki.fi/
Telephone: +358 (0) 2941 911

ISSN 1238-8645
ISBN 978-951-51-5336-4 (paperback)
ISBN 978-951-51-5337-1 (PDF)
Computing Reviews (1998) Classification: H.5.2, H.5.3, I.2.1
Helsinki 2019
Unigrafia

Human-Computer Co-Creativity — Designing, Evaluating and Modelling Computational Collaborators for Poetry Writing

Anna Kantosalo

Department of Computer Science
P.O. Box 68, FI-00014 University of Helsinki, Finland
anna.kantosalo@helsinki.fi

PhD Thesis, Series of Publications A, Report A-2019-3
Helsinki, August 2019, 74+86 pages
ISSN 1238-8645
ISBN 978-951-51-5336-4 (paperback)
ISBN 978-951-51-5337-1 (PDF)

Abstract

Human-computer co-creativity examines creative collaboration between humans and artificially intelligent computational agents. Human-computer co-creativity researchers assume that instead of using computational systems to merely automate creative tasks, computational creativity methods can be leveraged to design computational collaborators capable of sharing creative responsibility with a human collaborator. This has potential for extending both human and computational creative capability. This thesis focuses on the case of one human and one computational collaborator. More specifically this thesis studies how children collaborate with a computational collaborator called the Poetry Machine in the linguistically creative task of writing poems.

This thesis investigates three topics related to human-computer co-creativity: The design of human-computer co-creative systems, their evaluation and the modelling of human-computer co-creative processes. These topics are approached from two perspectives: an interaction design perspective and a computational creativity perspective. The interaction design perspective provides practical methods for the design and evaluation of interactive systems as well as methodological frameworks for analysing design practices in the field. The computational creativity perspective then again provides a theoretical view to the evaluation and modelling of human-computer co-creativity. The thesis itself consists of five papers.

This thesis starts with an analysis of the interaction design process for computational collaborators. The design process is examined through a review of case studies, and a thorough description of the design process of the Poetry Machine system described in Paper I. The review shows that several researchers in the field have assumed a user-centered design approach, but some good design practices, including the reporting of design decisions, iterative design and early testing with users are not yet fulfilled according to the best standards.

After illustrating the general design process, this thesis examines different approaches to the evaluation of human-computer co-creativity. Two case studies are conducted to evaluate the usability of and user experiences with the Poetry Machine system. The first evaluations are described in Paper II. They produced useful feedback for developing the system further. The second evaluation, described in Papers III and IV, investigates specific metrics for evaluating the co-creative writing experience in more detail. To promote the accumulation of design knowledge, special care is taken to report practical issues related to evaluating co-creative systems. These include, for example, issues related to formulating suitable evaluation tasks.

Finally the thesis considers modelling human-computer co-creativity. Paper V approaches modelling from a computationally creative perspective, by extending the creativity-as-a-search paradigm into co-creative systems. The new model highlights specific issues for interaction designers to be aware of when designing new computational collaborators.

Computing Reviews (1998) Categories and Subject Descriptors:
H.5.2 [User Interfaces]: User Interfaces — Evaluation
H.5.2 [User Interfaces]: Group and Organisation interfaces — Theory and models
I.2.1 [Artificial Intelligence]: Applications and Expert Systems

General Terms:
Design, Experimentation, Human Factors, Theory

Additional Key Words and Phrases:
thesis, computational creativity, human-computer co-creativity, co-creativity, human-computer interaction, evaluation, modelling, child-computer interaction, creativity support systems

Acknowledgements

Throughout the years, numerous people have had a positive impact on my academic thinking. I would like to thank you all for your support.

First of all, my supervisor, professor Hannu Toivonen has been an excellent guide and mentor on my academic journey. His call for a summer intern to work on interactive computationally creative writing piqued my interest and allowed me to combine my interests in human-computer interaction and classical science fiction. This internship grew into a Finnish Academy project and culminated in this PhD thesis, while professor Toivonen supported my personal growth in gradually giving me more responsibility within the project.

Special thanks are also in order for my mentor and co-author Dr Sirpa Riihiaho, who reviewed the evaluation plans for the Poetry Machine and conducted the second round of evaluations with me, and helped me to develop a more concise style of writing. A special mention also goes to all my other co-authors over the years, starting with Doctors Liisa Ilomäki, Minna Lakkala and Sami Paavola, who first introduced me to academic publishing; Dr Jukka Toivanen and Prof. Hannu Toivonen, who co-authored two papers with me on the design and evaluation of the Poetry Machine, with Dr Ping Xiao offering important feedback on the first; and once again to Prof. Toivonen, for our more theoretical work, which has since continued with Simo Linkola and Prof. Tomi Männistö. I am also grateful to my pre-examiners Prof. Georgios Yannakakis and Prof. Brian Magerko for their feedback on this thesis introduction.

In addition to my co-authors, several important discussions with various computational creativity professionals have affected my view of the field over the years. A great source of inspiration have been especially Dr Anna Jordanous, with whom I have discussed the nuances of computational creativity evaluation on many occasions; Prof. Geraint Wiggins, with whom I continue in math, where the English language fails; and Dr Oliver Bown, with whom I have shared interesting discussions on interaction design and computational creativity. I also want to thank the Discovery group, past and current members included, for nurturing an excellent, creative and friendly climate for academic growth.

On the practical side, this thesis would not have been possible without the skillful people, with whom I have worked on the CLiC project; special thanks go to Khalid Alnajjar, Mika Hämäläinen and Olli Alm. Likewise I am grateful for everyone who participated in collecting the data used in the related experiments, including Karoliina Tiuraniemi and Mikko Hynninen, who helped me with gathering the data for the first experiment; and Pekka Niklander, who graciously did not remark upon the very extended loan times of the equipment. Further my gratitude also goes to each and every teacher and pupil, who participated in the experiments illustrated in this thesis.

I am also grateful to several institutions for the financial support for this thesis, including the Doctoral School in Computer Science (personal grant), the Academy of Finland (CLiC and Algodan), the Department of Computer Science, the Helsinki Institute for Information Technology, and the European Commission (ConCreTe).

Finally I would like to thank some family and friends; my husband Jesse for his support on the English grammar and the many cups of tea brewed to fuel this writing process; my sister Emma and my dad Jarmo, who were an excellent child care alternative to the TV during the late nights I worked on my thesis manuscript; and my old history teacher Lottis for directing me towards the path of academic research.

Helsinki, June 27th, 2019
Anna Kantosalo

Contents

List of Publications

This thesis consists of five original papers referred to as Papers I-V in this thesis introduction. The papers are listed below, with a record of the contributions of each author. Reprints of Papers I, II and V, and accepted author manuscripts for papers III and IV are included at the end of this thesis introduction.

Paper I

KANTOSALO A., TOIVANEN, J. M., XIAO, P., AND TOIVONEN, H. From Isolation to Involvement: Adapting Machine Creativity Software to Support Human-Computer Co-Creation. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, 10-13, June, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., pp. 1-7

Author roles: I did the background research for this paper and wrote the main parts of it. Dr. Toivanen wrote the sections describing the poetry generation methods of the Poetry Machine system. Dr. Toivanen and Prof. Toivonen participated in the design of the Poetry Machine. The description of the Poetry Machine through Wiggins' Creative Systems Framework is based on discussions between me, Dr. Toivanen and Prof. Toivonen. Dr. Xiao and Prof. Toivonen provided important remarks and questions for improving the original manuscript for the target audience.

Paper II

KANTOSALO, A., TOIVANEN, J. M., AND TOIVONEN, H. Interaction Evaluation for Human-Computer Co-Creativity. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, USA, June 29 - July 2, 2015), H. Toivonen, S. Colton, M. Cook, and D. Ventura, Eds., pp. 274-283

Author roles: I conducted the evaluations and background research described in this paper. Dr. Toivanen participated in the design of the Poetry Machine and provided the algorithms and their description for the article. Prof. Toivonen discussed the manuscript with me and suggested some changes.

Paper III

KANTOSALO, A. AND RIIHIAHO, S. Experience evaluations for human-computer co-creative processes — planning and conducting an evaluation in practice. *Connection Science* 31, 1 (2019), pp. 60-81

Author roles: I suggested the evaluation strategy presented in this paper and discussed it with Dr. Riihiaho. We conducted the illustrated experiments together and did the analysis together. I wrote the first draft of the manuscript and we revised it collaboratively.

Paper IV

KANTOSALO, A. AND RIIHIAHO, S. Quantifying co-creative writing experiences. *Digital Creativity* 30, 1, (2019), pp. 23-38

Author roles: I suggested the evaluation strategy presented in this paper and discussed it with Dr. Riihiaho. We conducted ten of the illustrated experiments together, while Dr. Riihiaho conducted two alone. We contributed equally to the manuscript and the authors are listed in alphabetical order.

Paper V

KANTOSALO, A. AND TOIVONEN, H. Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In *Proceedings of the Seventh International Conference on Computational Creativity* (Paris, France, 27 June - 1 July, 2016), F. Pachet, A. Cardoso, V. Corruble, and F. Ghedini, Eds., pp. 77-84

Author roles: The article is based on the discussions I had with Prof. Toivonen on Wiggins' Creative Systems Framework and its applications to co-creativity. Paper V continues the work started in paper I by formalising the framework in a way that could help in the design of co-creative tools. I wrote the main part of the Introduction and the rest of the paper was written collaboratively based on discussions and drafts made together with Prof. Toivonen.

Chapter 1

Introduction

Increasing human potential is the underlying incentive for all technology. In creativity, technology can be used to facilitate faster creation, to improve human creative capability through training, and to enable completely new ways to create [89]. Interactive computational creative systems can already prompt users to think of ideas they would not have thought of otherwise [6]. With the improvement of computational creativity theory and methods, it is possible to go beyond using software as a creative tool, and invent systems with which humans can collaborate analogous to human creative partners.

Human-computer co-creativity studies how to facilitate human creativity via computationally creative means and vice versa. It goes beyond the traditional goals of human creativity support, by considering the creative interplay between humans and artificially intelligent agents instead of merely the use of digital creativity support tools such as text editors. By combining their abilities, the human and the creative computer can achieve better outcomes together than either would have achieved alone.

To leverage the full potential of human-computer co-creativity, it needs to be studied as a complex phenomenon. This thesis utilises methods from interaction design and computational creativity theory to study different aspects of human-computer co-creativity. It illustrates the design process of human-computer co-creative partnerships, their evaluation, and how to model the creative process between a human and a computer by using computational creativity models. The creative domain of the work is linguistic creativity, more specifically, poetry. The majority of work in this thesis was conducted as case studies with a computationally creative system called Poetry Machine. This introduction to the thesis contextualises the findings of the case studies in a frame of reference illustrating their relationship to emerging trends in human-computer co-creativity research.

1.1 Research Questions

In this thesis, I examine human-computer co-creativity from three perspectives: design, evaluation, and modelling human-computer co-creativity. I have defined four research questions related to these topics: one for design, two for evaluation, and one for modelling.

The first topic, designing human-computer co-creativity, is investigated in relation to interaction design. I examine the development process of the Poetry Machine tool as an interaction design project, and compare it and other examples to the standard of user-centered design. My research question is: *How does the design process of human-computer co-creative systems differ from typical design processes of interactive systems?*

The second topic, evaluation of human-computer co-creativity, is divided into qualitative and quantitative evaluations conducted with two versions of the Poetry Machine prototype. The second question asks, *how can qualitative evaluation guide the design of a co-creative system?* The third investigates, *how can quantitative evaluation be used to compare different co-creative processes in a meaningful way?*

The final topic, modelling human-computer co-creativity focuses on the co-creative process as viewed from a computational perspective. Modelling the creative process is important in order to better understand what happens in the creative process between a human and a computer. More specifically, I ask, *how can the human-computer co-creative process be described in a way that can be used to guide design decisions?*

1.2 Methodological Frameworks

This thesis combines both a theoretical and an experimental approach to human-computer co-creativity research. The theoretical aspects of human-computer co-creativity are studied through a lens based on computational creativity research, while the experimental methodology is provided by the field of interaction design. I will consult computational creativity literature especially in modelling human-computer co-creativity, but theoretical aspects of the evaluation of creativity and defining human-computer co-creativity draw also from the body of work done in the field of computational creativity. Interaction design methods form the backbone for the practical design and evaluation of the Poetry Machine.

1.2.1 Computational Creativity

Computational Creativity is a sub-field of artificial intelligence focused on the study of creative systems via computational means. The field encompasses a variety of domains, including autonomous creativity, simulation of creativity, and human-computer co-creativity. Although the idea of machine creativity was already discussed by the pioneers of artificial intelligence, computational creativity itself is still an emerging field. Systematic scientific work on computational creativity started in the mid-1990s [18].

Theories of computational creativity offer a way to view human-computer co-creativity as a process combining human and computational elements. This allows not only the categorisation of current computational systems based on their role in the creative process, but also the consideration of factors enabling these systems to take a more balanced role in future co-creative processes with humans. This thesis draws especially on Wiggins' Creative Systems Framework [124, 125] to define a model for human-computer co-creativity. Our model, presented in Paper V, shows one way of categorising the roles of human and computational collaborators in the co-creative process.

Computational creativity definitions can also be used to separate computational creativity from traditional creativity support tools: Colton and Wiggins consider that computationally creative systems take on some responsibilities through which they exhibit creative behaviour [25], whereas creativity support systems do not. We have echoed this idea in our definition of human-computer co-creativity in Paper I.

1.2.2 Interaction Design

Interaction design is the holistic study of the relationship between designed artefacts, those that are exposed to them, and the socio-cultural context surrounding their relationships [43]. It is an umbrella term for different fields considering human interaction with an environment and designed objects, including human-computer interaction, human factors and ergonomics, user experience, and anthropology [102, pp. 9-10]. Unlike these related academic fields, interaction design actively seeks to improve interactive systems via participatory design practice [43].

Interaction design has gradually gained momentum as a go-to methodology for investigating human-computer co-creativity in practice. In the same conference in which our paper I was published, Bown [8] suggested interaction design as a way to ground empirical evaluations of computational creativity. Later Bown [10] and Yee-King and d'Inverno [130] also argued for a stronger focus on the experiences of humans interacting with creative systems, suggesting a need for further integration of interaction design practice into human-computer co-creativity research. Lately more traditional interaction design tools have been adopted to the design of human-computer co-creative systems. Examples are discussed in Chapter 3.

I originally adopted interaction design as the methodological framework for this thesis since it enables the study of human-computer co-creativity in situ, with real applications and users. This follows Fallman's [43] characterisation of interaction design as three overlapping activities: *design practice, design exploration and design studies.* Through participating in the design process via *design practice* interaction designers can prepare design artefacts, which can be used to investigate research questions in context via *design exploration.* And through *design studies* the practice of interaction design itself can be improved by incorporating findings from design practice and exploration. Following this framework, this thesis itself forms a part of design studies reporting on the design practice and exploration in human-computer co-creativity through the design and evaluation of the Poetry Machine system.

1.3 Research Context

Specifying a research context is extremely important for interaction design. It affects everything from what is viable to design to the methods used in the project. The case studies outlined in this thesis all investigate human-computer co-creativity through a co-creative tool called the Poetry Machine. The intended user group of the Poetry Machine is children and the application itself functions within the domain of computational linguistic creativity.

1.3.1 Children as Users

The Poetry Machine is designed to be used by children at school. The case studies discussed in this thesis have all been conducted with 9-11-year-old children. All participants spoke good Finnish and evaluated the Finnish language version of the Poetry Machine.

Working with children poses some specific constraints for interaction design. As a sensitive user group children require specific ethical considerations, including acquiring informed consent from their guardians. They also require evaluation methods that take into account their developing skills (see e.g. [41, 85, 116]). All of the evaluations in this thesis were conducted with children working in pairs to make testing more comfortable for them. The tests were also more limited in length, as lengthy surveys may cause bias when children attempt to satisfy requirements instead of answering according to their experiences and feelings [95]. Further details on the characteristics of children as users can be found in a review on interaction design and children by Hourcade [59].

1.3.2 The Poetry Machine

As a design artefact, the Poetry Machine has undergone major changes during the work conducted for this thesis, but its functionality has remained very similar through the two main versions. The first prototype of the Poetry Machine tool was constructed by myself together with Jukka Toivanen and Hannu Toivonen, with input from the target user group and pedagogical researchers Liisa Ilomäki and Minna Lakkala. Jukka Toivanen developed the poetry generation methods for the first prototype based on his previous work in the area of computational linguistic creativity (see methods in [113, 114]). This version was used in the evaluations reported in Paper II. Based on these evaluations the Poetry Machine system was developed further with Mika Hämäläinen, Olli Alm and Khalid Alnajjar working in the development team, Sari Laakso and Minna Lakkala as usability consultants, and Hannamari Vahtikari as a graphic designer. The current version of the system uses poetry generation methods developed by Mika Hämäläinen [53]. The details of the generation methodologies used by different versions of the Poetry Machine tool are outside the scope of this thesis.

In a typical use case with the Poetry Machine, the user starts by selecting a topic from a list of child-friendly themes, including for example family, seasons, and vehicles. Alternatively the user can select "random" and have the Poetry Machine select one of the themes at random. Based on this prompt, the Poetry Machine generates a small poem excerpt (the first version provided excerpts of varying length, the second provides five lines). The user can then modify this excerpt via a drag-and-drop based interface. The interface itself was inspired by fridge magnet poetry, which allows anyone to compose a poem by rearranging a set of magnetic word tags on a metallic surface. The Poetry Machine also allows users to remove words or rows entirely and add their own words or lines. The users can also prompt the Poetry Machine for more material, including rhymes and alliterations for specific words, substitute words for a specific word in context, or new lines for the poem.

The linguistic domain has affected this thesis in many ways. The influence has been most direct in designing interactions with the Poetry Machine system. One of the important findings of this thesis is that the chosen interaction modalities and the planned role of the system have in turn affected the development of the poetry generation methods used by the system.

1.4 Structure

This thesis consists of five papers related to the topics of design, evaluation, and modelling of human-computer co-creativity. Papers I–IV describe case studies conducted with the Poetry Machine tool and children. Paper V has a more theoretical note. This thesis introduction contextualises the findings of the papers, illustrating their relationship to emerging trends in human-computer co-creativity.

This thesis introduction continues with a brief background to human-computer co-creativity illustrating a short history and related terminology. It then moves on to the theme of designing human-computer co-creativity. Chapter 3 accompanies Paper I. Together they focus on my first research question, *how does the design process of human-computer co-creative systems differ from typical design processes of interactive systems?* Chapter 4 discusses the evaluation of human-computer co-creativity. Paper II focuses on my second research question, *how can qualitative evaluation guide the design of a co-creative system?* And Papers III and IV consider my third research question, *how can quantitative evaluation be used to compare different co-creative processes in a meaningful way?* Chapter 5 considers the modelling of human-computer co-creative processes. Together with Paper V, they answer my final research question, *how can the human-computer co-creative process be described in a way that can be used to guide design decisions?* This thesis introduction ends with a conclusion, summarising my findings on each research question and ideas about future work.

Chapter 2

Human-Computer Co-Creativity

Human-Computer Co-Creativity refers to the collaboration between a human and an artificial intelligence system on a creative task. In literature, multiple terms are used to describe roughly the same ideas related to sharing and distributing co-creative responsibility between a human and a computational author. Candy and Edmonds [17] call collaborative creativity between humans, computers, or both simply co-creativity. Other terms used in literature include mixed-initiative co-creativity [128] and mixed-initiative creative interfaces [34], which both emphasize the computational system's capacity to initiate interaction resulting in creative outputs. In this thesis, I use the term human-computer co-creativity, as it fits different variations of creative collaboration, but requires the involvement of at least one human and one computational agent. This thesis focuses on the case of exactly one human and one computational collaborator.

This background chapter defines the terminology needed for discussing Human-Computer Co-Creativity. I start by a brief discussion of the history of the field, focusing on the groundwork laid by research in Creativity Support Systems and Interactive Computational Creativity systems. I then proceed to present a working definition for Human-Computer Co-Creativity based on Rhodes' [97] 4P's of creativity framework for the purposes of this thesis.

Since the beginning of the thesis project, many case studies have been published in different domains of human-computer co-creativity. In this thesis introduction, I focus on work done in linguistic creativity contexts. I only cite work conducted in other domains when it considers aspects of human-computer co-creativity and interaction design from a more general perspective.

2.1 Origins of Human-Computer Co-Creativity

In this section I briefly consider the history of human-computer co-creativity, and give an introduction to its related fields: Creativity Support Tools, and Computational Creativity.

2.1.1 History of Human-Computer Co-Creativity

The idea of creative computers was already discussed by many early pioneers of computing such as Turing and Shannon [25]. The idea of a machine partnering with humans in solving hard problems also emerged during the early years of artificial intelligence research [5]. In 1960 Licklider [77] famously wrote of a man-computer symbiosis. He expected machines to facilitate formulative thinking, take over routine work and cooperate with humans in decision making going beyond their predetermined programming. In his vision, computers would ultimately outdo humans in thinking, but the man-machine symbiosis would be an unavoidable phase in the development of the autonomous machines, during which humanity would enjoy unprecedented intellectual creativity.

Today creative computers are one of the key foci of computational creativity research and the man-computer symbiosis is facilitated by interaction design, which in turn has segmented into further subfields such as creativity support systems. Human-computer co-creativity has emerged from combining computational creativity and creativity support system research [29, 31]. Current research on human-computer co-creativity can be viewed on a spectrum between these two fields. Detering et al. [34] consider that the ends of this spectrum represent different initiatives: computational creativity focuses on the computational initiative, while creativity support systems focus on the human initiative.

In my view, current examples of human-computer co-creative systems can be examined from three perspectives: a human-computer co-creative perspective, a computational creativity perspective, and a creativity support systems perspective. The perspective varies according to the focus of research, which can be on the human-computer collective or on one of the collaborators. For example, information flow can be considered from all three perspectives: the human-computer co-creativity perspective focuses on what type of information needs to be exchanged to best facilitate co-creativity, while the computational creativity perspective focuses on how this information is processed and produced by the computational collaborator, and the creativity support perspective looks at how the human would like to receive this and input similar information.

Both computational creativity and creativity support systems are young fields of research themselves: Systematic research into computational creativity has been carried out since the turn of the millennium [25]. Efforts to establish the field of creativity support systems started around the same time [40, 44]. It did not take long for the two research paradigms to start mixing, as researchers in both fields started to consider the benefits of merging efforts. Many projects in computational creativity began to include interaction design and in 2009 Morris and Secretan [87] suggested that computational creativity methods could be leveraged to create better creativity support systems.

As the domain of human-computer co-creativity started to gradually take shape, a number of definitions for human-computer co-creativity were suggested close to each other: In 2013 Davis [29] proposed human-computer co-creativity as a way of enabling computers to contribute as a partner in the creative process. In the following year Yannakakis et al. [128] defined mixed-initiative co-creation as “the task of creating artifacts via the interaction of a human initiative and a computational initiative”. In the same year we defined Human-Computer Co-Creativity as collaborative creativity characterised by a shared responsibility between the human and the computational participant over the created artefact (see paper I). While the definitions for the term human-computer co-creativity continue to evolve, current research covers different styles of computational collaborators. These include task-divided co-creative systems with clearly defined tasks and responsibilities defined in our Paper V as well as computational colleagues capable of taking the initiative based on the limited self-awareness described in [30].

2.1.2 Creativity Support Tools

Creativity support tools are a multidisciplinary research field combining computer science, psychology, human-computer interaction, information systems, information visualisation and software engineering [106]. Instruments used in creativity support tool research are typically drawn from interaction design and thus creativity support tool literature offers examples of applying interaction design methods to creative domains.

Any tool that can be used in the open-ended creation of new artefacts is a creativity support tool [21]. Examples range from individual creativity support tools, such as video editing software, to collaborative creativity support tools, such as sharing the videos on popular platforms [105]. Creativity support tools can also be combined into larger, at times also physical, creativity support environments [21]. It is sometimes difficult to make a distinction between a productivity support tool and a creativity support tool, as their definition depends on the task: For example, word processors can be used for both routine work and creative writing [21].

Creativity itself can be supported in many ways. Lubart [80] considers four categories for promoting human creativity with computational means: managing creative work, facilitating communication between individuals, suggesting creativity enhancement techniques, and human-computer co-operation. Creativity support tool research can also help invent completely new methods or domains of creativity [89, 105]. Human-computer co-creative systems can also be used for these tasks, but the full power of their artificial intelligence methods is only leveraged in more complex tasks involving human-computer co-operation.

Two recent reviews of case studies give some insight into what types of systems have been studied in creativity support tool research. Gabriel et al. [47] analysed 49 creativity support tools and Wang and Nickerson [120] reviewed 48 individual creativity support tools including both general and domain-specific tools (7 of the tools were reviewed in both studies). Based on the studies, ideation [47, 120] and evaluation [47] are the most supported tasks in current systems. Most of the tools evaluated by Gabriel et al. also supported either remote or co-located collaboration between humans. The collaborative tools were typically used via interactive tabletops or whiteboards, while around a half of the individual systems in the study were mainly used via web-based interfaces.

2.1.3 Interactive Computational Creativity

Computational creativity research includes both theory and practise. While the theoretical side of computational creativity considers topics such as simulation of human creativity or the evaluation of computational creativity, the practical side typically attempts to find methods for generating creative artefacts in a particular domain. These generation methods are often exemplified in autonomous agents or interactive systems. For the purposes of this thesis, it is useful to look at how interactive computational creativity systems can be categorised, and how the Poetry Machine fits in.

On a general level, systems within the field of computational creativity can be categorised by the domain of creativity they work in. These include both traditional creative fields, such as music, art, and literature and less traditional fields, such as choreography, cooking recipes, and humour [79]. The Poetry Machine works within the linguistic creativity domain.

Pérez y Pérez [94] suggests that computational creativity research exists in a spectrum between two major paradigms: an engineering-mathematical oriented and a cognitive-social oriented approach. The selected paradigm affects what is studied and what kind of methods are used: The engineering-mathematical approach focuses on optimisation techniques and patterns, while the cognitive-social approach focuses on simulating human creative processes and proofing theories of human creative cognition with computational means. Researchers often adopt methods from one paradigm to the other, using them as a sort of supporting infrastructure to study questions specific to the researchers' own research paradigm. This causes difficulties when researchers from opposite paradigms attempt to understand each other, evaluating the other's research against a different set of relevant questions. Pérez y Pérez calls this the "tower of Babel effect". In this thesis, the Poetry Machine is approached from the cognitive-social perspective, while its computational creativity methods remain in an instrumental role.

In addition to these general categories, interactive computational creativity systems can also be categorised by the role and number of creators: Maher [82] investigated creative ideation, and considered who is being creative when a human uses an interactive computational creativity system. In her sample of early systems, humans could *model* the computational generation methods or processes, or *generate* artefacts assisted by the system. The computational system could then *support* the generative act, *enhance* the abilities of the human or *generate* artefacts. Viewed through Maher's framework, both the Poetry Machine and the human user generate parts of the creative artefact produced in the co-creative process.

Maher [82] also categorised interaction between humans and computational creativity systems by the number of humans and systems participating in it. In her categorisation both humans and computational systems can participate *individually*, or in small *groups*, which she calls *teams* in the case of computational systems. On the other hand, people can also participate as a *society*, represented for example by crowd sourcing. Correspondingly, computational agents can participate as *multi-agent societies* with distributed control. The Poetry Machine participates in interactions as an individual, while our experiments conducted with it include both individual and pairs of humans (see papers II, III and IV).

Maher's [82] categorisations also seem useful for describing human-computer co-creativity: The computer roles support, enhance and generate seem to describe a gradual shift from creativity support systems towards more co-creative systems, while the human roles of defining the computational models or generating with the system help to deduce some finer nuances between computational creativity and human-computer co-creativity. For example, Maher describes the highly autonomous Painting Fool [23] as a generative system with the human as a model developer.

Classifying computationally creative systems can however be difficult because of the lack of data. For example, Maher [82] mentions the DARCI system [93] as an example of both human and computational generative creativity. However, later publications of the DARCI system clearly demonstrate it is intended as an autonomous system, generating art without humans (see e.g. [92]). Conversely Maher defines the Curious Whispers project as an example of the human as model developer, although a later description of the system shows humans creating together with the system as generators [103]. It therefore appears that promising work done in interactive creativity may in time shift towards co-creativity, autonomous computational creativity, or creativity support systems, and when novel systems in the field are reported, their descriptions often lack enough detail to fully recognise their purpose, limitations, and potential.

2.2 4P's of Human-Computer Co-Cretivity

In paper I, we defined human-computer co-creativity as "collaborative creativity where both the human and the computer take creative responsibility for the generation of a creative artefact". In order for this definition to be useful *collaboration* and *creativity* need to be defined in a meaningful way.

Terveen [112] defines *human-computer collaboration* in the context of problem solving as a process in which at least one human and one computational agent work together to achieve shared goals. This requires agreeing on the goals, planning, allocating responsibility, coordination, sharing context, communication, adaptation, and learning. Most of these requirements are also applicable to creative collaboration, although their importance depends on the context and the participating individuals.

Defining *creativity* is more difficult. Literature has considered the domain, extent, or underlying cause of creativity [67], but a good definition is difficult to find, as creativity research itself is segmented into subfields [55]. Rhodes' [97] classical 4P's of creativity have been used in computational creativity literature to examine creativity from different angles (see e.g. [27, 68, 76]). They offer four perspectives to creativity:

- **Person**: The active creative individual.
- **Process**: The process through which creativity is manifested.
- **Product**: The end result of a creative process.
- **Press**: The environment and history of the creative individual.

Together the 4P's form an interconnected description of creativity: The creative person participates in the creative process, generating a product in a constant exchange with the creative press. By manipulating one perspective other perspectives can also be changed. By understanding each perspective we will gain a more thorough understanding of a system and how it can be improved to enhance creativity.

These perspectives offer a useful way for decomposing human-computer co-creativity. Jordanous [68] already extended the framework for computational creativity, dubbing the person as producer to allow for computational agents. In this thesis introduction I use the person perspective to consider both participants of the human-computer collaboration, referred to individually as the human and the computational collaborator, and together as a collective. In the next sections I use the process perspective to discuss the interactions within the creative collective, while the individual processes of the collaborators are viewed as demonstrations of their skills. I use the press perspective to consider the context and reception of human-computer co-creativity and the product perspective to discuss the attribution of creative responsibility within a collective.

2.3 The Creative Producers

The human-computer co-creative setting in this thesis involves two creative producers. I refer to them as the human collaborator and the computational collaborator.

2.3.1 The Human Collaborator

Individual traits of the human collaborator have been a strong focus of both contemporary and past research in human creativity [55, 68]. Factors that affect co-creativity between humans may also be relevant in designing human-computer co-creativity. These include individual qualities, such as task motivation, domain knowledge and creative thinking skills [2], and interpersonal qualities, including how well two creative partners complement each other, interpersonal facility, gender and age [1]. Collaboration itself may also affect these individual qualities [1].

In this thesis I approach the human collaborator from an interaction design perspective as an example member of a user group. Typical user traits considered in interaction design include knowledge, skill, experience, education, training, physical attributes, habits, preferences and other capabilities [62]. As part of the design process of the Poetry Machine tool, I investigated the general requirements of the user group of Finnish primary school children and observed future users in situ at a school. This stage of the design process is described in paper I, and incorporation of the characteristics of users into evaluation is discussed in papers II, and III.

2.3.2 The Computational Collaborator

According to Jordanous [68] some attempts have been made to describe the qualities of computationally creative systems by imitating the approach of describing traits of a creative human. Such models include, for example, Colton's Creative Tripod [24], which requires systems to demonstrate skill, appreciation and imagination to be considered creative.

Jordanous notes that computational creativity systems typically work within one domain and systems tend to be built around specific skills required within it. These skills can be seen to be embodied in the algorithms and knowledge bases of the system. Therefore the skill and capacity of many current computational collaborators can be effectively described by considering the computational creativity methodology used in them; The Poetry Machine uses a set of heuristics to generate new poems using human authored example structures and word databases. In addition to pre-defined heuristic structures [22, 119] possible linguistic co-creativity skills include for example neural language models [22, 73, 101], probabilistic models [121], and case-based reasoning [100].

Computational creativity researchers often strive to construct autonomously creative systems. Creative autonomy entails the ability to independently apply and change the standards used in generating and evaluating creative products [65]. However, autonomy is not a necessary requirement for computational collaborators. In Paper I we investigated the transformation of autonomously evaluative computational creativity algorithms into co-creative systems and found that in many cases the human collaborator's role needed to be increased in the creative process in order to meaningfully interact with the system.

Our findings could be seen as decreasing the autonomy of the computational collaborator. However, creative autonomy and co-creativity are not contradictory to each other either: Creative autonomy can be achieved by intentional, non-random changes to the generation and evaluation methods used by the computational agent [65], facilitated by self-awareness over their status [78]. Intentionality and limited self-awareness have been suggested as qualities for computational collaborators participating in human-computer co-creativity [30], and an example system, the Drawing Apprentice [31], has successfully operationalised these concepts in co-creative drawing. Therefore it seems that autonomy and related concepts are useful descriptors for some computational collaborators, but meaningful co-creative experiences can be supported also by collaborators without autonomous capabilities.

The basis of creative collaboration lies in communication. Communication is affected by multiple factors, such as how the communication happens, what is communicated, and how the communicators are represented. On a primary level communication in human-computer co-creativity is affected by the input and output mechanisms available to the collaborators. In multimodal interaction the collaborators can use multiple channels to communicate with each other. A number of different channels are available, including channels corresponding roughly to human senses, including visual, aural, haptic, gustatory and olfactory as well as combinatorial channels such as keyboard, mouse and motions [64]. To my knowledge, no survey has been carried out to investigate the effect of specific communication channels to co-creativity, but co-creative applications include examples utilising various inputs and outputs. The Poetry Machine communicates to the human visually, combining graphical elements and text, while the human can communicate to the system with specific operations available through a point and click interface and input text through a keyboard.

In addition to communication channels, it is also important to consider what is communicated. Many computational collaborators, including the Poetry Machine, mainly collaborate by sharing creative artefacts with the human collaborator. However, in co-creativity between humans, discussion about ideas and the communication of affect are important for success [3]. Communicating affect or discussing ideas would require meta-level processing from the computational collaborator. Although affective computing has been studied elsewhere, first attempts to use it in human-computer co-creativity development have only started to emerge. These include for example the LuminAI system [126], which assesses the emotion of its human collaborator in order to decide what role it should assume in dance improvisation.

The representation of the computational collaborator is also important. Should the computational collaborator emulate human appearance and to what extent? Is the system embodied, or used in an application running in a device? Many current co-creativity systems, including the Poetry Machine, operate as non-anthropomorphic web applications. A few recent systems use embodied interactions, such as an interview-dialogue story-telling system build for the Nao robot [123] and an interactive task demonstration algorithm for robots [45]. Humans have a tendency to get attached to anthropomorphic systems, project normative traits to them based on their assumed gender, and prefer to interact with machines with specific personalities [132]. Therefore it is reasonable to assume that embodied or anthropomorphised co-creative systems will be at least partially subject to the same considerations. Further research is needed to investigate how these factors affect human-computer co-creativity.

2.4 The Creative Process

Understanding the creative process is important in order to motivate and teach it [97], to measure it [88] and to improve it [55, 88]. Here I focus on the joint process between the human and the computer, what factors may influence it and how the roles of the human and the computational collaborator manifest in it. I view the process as an exchange of creative artefacts, evaluations, and meta-information that facilitates communication. This exchange happens between two or more participants using different communication methods. Each collaborator does not necessarily produce the same information, or use the same information as the others. Modelling the human-computer co-creative process is discussed in Chapter 5.

2.4.1 Dichotomies of the Co-Creative Process

Co-creative processes can be categorised in different ways. Abra [1] suggests four dichotomies to describe co-creative processes: fixed vs. on-going, intimate vs. remote, horizontal vs. hierarchical, and homogenous vs. heterogeneous. These can also be applied to the human-computer co-creative process.

The *fixed vs. on-going* dichotomy deals with time: does the process have a fixed deadline or does it extend over longer time frames. Examples in human-computer co-creativity literature typically focus on simple experiments conducted in laboratory environments, describing human-computer co-creativity in a fixed time frame. For example, all experiments described in this thesis were conducted in a short period of time. A few examples, especially from musical co-creativity, describe computational collaborators which have been in a long-term collaboration with their human collaborators (see e.g. [7, 16]).

The *intimate vs. remote* dichotomy refers to whether the collaborators are co-located or not. As both computational collaborators and remote human collaborators are limited in their communication modalities, it could be argued that human-computer co-creativity resembles remote co-creativity between humans. However, collaboration with physical computational collaborators may be more similar to intimate collaboration with humans.

The *horizontal vs. hierarchical* dichotomy considers the organisation of the creative process. Horizontal collaboration implies equal decision-making power between the collaborators, while hierarchical collaboration introduces dominance and power considerations into their relationship. Many human-computer co-creative systems support a hierarchical relationship in which the human has authority over the computational collaborator. This is visible in systems that give priority to the human collaborator's inputs (e.g. the Tanagra [108]), or allow limited manipulation of the creative product without human intervention (e.g. the Poetry Machine). D'Inverno and McCormack [36] suggest that most artists prefer a hierarchical relationship when working with artificial intelligence — the system should serve the artist's goals, while the artist claims the honour.

The *homogenous vs. heterogeneous* dichotomy refers to the distribution of different tasks among the collaborators. In homogenous collaboration the collaborators work on similar tasks, while in heterogeneous collaboration they can focus on different tasks. Various hierarchies exist for distributing the heterogeneous work effort in a creative process between humans [84, 91]. The idea of homogenous vs. heterogenous collaboration has also been suggested for human-computer co-creativity in the work by Yannakakis et al. [128] and in our task divided co-creation model described in paper V and discussed in Chapter 5.

2.4.2 Roles

Several roles have been suggested for computers in creative processes. As shown in Table 2.1 the roles can be grouped into four categories: Support, enhance, collaborate and other.

The support and enhance categories in Table 2.1 receive their name from Maher's [82] classification. Computers in the support role provide humans with tools and techniques to support human creativity [82]. This includes several instrumental roles, such as nanny, pen-pal [80], environment, or toolkit [90], and roles for training creativity, such as dumbbells and coach [89]. Creativity training can also be considered a part of Maher's enhancement role, which focuses on computational systems extending the abilities of humans by presenting information or enhancing creative cognition. This role also includes Nakakoji's [89] class of running shoes, which describes systems intended to enable faster creation for humans.

Several, more specific roles for the computational collaborator are defined within the collaborate category. It includes Lubart's [80] computa-

<table>
<tr><th></th><th>Maher, 2012 [82]</th><th>Negrete-Yankelevich & Morales-Zaragoza, 2014 [90]</th><th>Kantosalo & Toivonen, 2016 (Paper V)</th><th>Nakakoji, 2006 [89]</th><th>Lubart, 2005 [80]</th></tr>
<tr><td rowspan="2">Support</td><td rowspan="2">Support</td><td rowspan="2">Environment
Toolkit</td><td rowspan="2"></td><td></td><td>Nanny
Pen-pal</td></tr>
<tr><td rowspan="2">Dumbbells</td><td rowspan="2">Coach</td></tr>
<tr><td rowspan="2">Enhance</td><td rowspan="2">Enhance</td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td>Running -
Shoes</td><td></td></tr>
<tr><td>Collaborate</td><td>Generator</td><td>Generator
Apprentice
Master</td><td>Generator
Evaluator
Definer</td><td></td><td>Colleague</td></tr>
<tr><td>Other</td><td></td><td></td><td>Pleasing
Provoking</td><td>Skis</td><td></td></tr>
</table>

Table 2.1: Roles for computers in the creative process.

tional colleagues, which represent to him the ultimate creativity support tool. Our model of task-divided creativity, defined in Paper V, suggests the computational collaborator can act as a concept generator generating artefacts fitting a specific conceptual description, evaluate these concepts as concept evaluator, or define the conceptual space itself as a concept definer. The generator class is also acknowledged by Maher [82] and Negrete-Yankelevich and Morales-Zaragoza [90]. Negrete-Yankelevich and Morales-Zaragoza define an additive model in which the first generative collaborator can be improved on by adding more complex evaluative roles. In the apprentice role systems still require human evaluative intervention, but systems in the master role are capable of conducting full evaluations on their own, although the human still configures the system and thus decides the conceptual space.

The final category, other, includes roles that can be seen as complementary to the basic roles suggested by the three other categories. This includes our pleasing and provoking agents, which describe collaborator responses to human input (Paper V), and Nakakoji's [89] role of skis, which entails the capability of computational tools to leverage completely new creative domains. Finally a number of additional domain-dependent roles have been suggested for computational collaborators (see e.g. [12, 28]). A discussion of domain-dependent roles is outside the scope of this thesis.

Possible roles for the human collaborator include modelling and generating [82], which extend the role of the human towards the programmer of the program. The human's role can also reflect the roles suggested for the computational colleague, including for example different types of assistantship or partnership [72]. The roles defined for task-divided co-creativity in our model (Paper V) can also be used to describe the roles of human collaborators.

2.5 The Creative Press

The creative press refers to the context and the environment in which the creative process happens. This is not limited to the immediate physical and intellectual climate surrounding the creative activity. The creative press also includes the continuum of influences an individual gathers throughout their life [97]. In computational creativity the press includes the context of the computationally creative entity, interactions with it and audience bias towards it [68]. As the previous sub-section focused on the interaction as part of the creative process, in this subsection I focus on the context and the reception of human-computer co-creativity among human collaborators as well as in society.

2.5.1 Context

Following interaction design practises, the context of human-computer co-creativity is typically represented as the context of use. Traditionally the context of use has been fairly static, and most human-computer co-creative projects focus on specific use cases, such as interior design [37], computer game level design [108], or sketching [32]. At times the environment of use is also defined to be, for example, an office [37]. The Poetry Machine is designed for poetry generation in a specific environment, the school.

A subfield of interaction design, context-aware computing, focuses on delivering experiences for changing contexts. In context-aware computing, the context is understood as information that can be used to characterise the situation of different entities, such as persons, objects, and places [35]. Understanding the context as a changing element is gradually becoming a success factor also in human-computer co-creativity. Current non-context-aware computational collaborators may for example make linguistic suggestions that are completely opposite to a user's intended message [22].

A context-aware computational collaborator could produce more relevant and useful material corresponding to the mental state of the human collaborator. The Digital Improv Project in interactive narrative is currently investigating how a shared mental model between the collaborators could be negotiated in practise [57, 58].

The context, whether understood as a static or a changing element of co-creation, should influence the specific interaction mechanisms and computational generation methods used to design a useful and meaningful computational collaborator. It appears that there are specific interaction methodologies that are more useful for a specific context, even within the same domain: Clark et al. [22] studied two linguistic co-creativity systems and found that the users of their story-writing system enjoyed complete sentence suggestions, while users of their slogan-writing system reported a need for single word suggestions. Interaction designers designing computational collaborators should therefore investigate the context of use carefully and utilise it in both design and evaluation of their systems. Paper I outlines how the school context was taken into account in designing the Poetry Machine, while Papers II and III consider how the context affected the evaluation of the Poetry Machine.

2.5.2 Reception

Humans assess the same co-creative system differently when they themselves collaborate with it and when they observe others collaborating with it [10]. The reception of human-computer co-creativity can be viewed from these two perspectives: from within the collective, and outside of it. Considering the societal and individual response to human-computer co-creativity is important, as the press perspective is often entangled with other perspectives [76]. We can also consider the reception of human-computer co-creativity from the other perspectives provided by the 4P's model, including the reception of the computational collaborator and its products, or the collective process and the collective product.

A key element in an individual human collaborator's response to a creative system is the user experience provided by the system. Brown [16] suggests that interaction design can be used to heighten the sense of a genuine partnership and a sense of agency. At best, interaction with even a relatively simple computational collaborator can feel like interaction with humans, culminating in experiences of productive collaboration and eureka moments [22]. It also appears that the personal characteristics of the human collaborators influence their responses, for example, novice writers seem to be more keen to accept computational collaborators in their own creative processes than professionals [22]. The artist's willingness to adopt a computational collaborator into the creative process may also be affected by other societal considerations in their immediate surroundings. Colton [24] notes that artists using computers in any fashion may be shunned in their own fields.

The response to the contributions of the computational collaborator also varies across human collaborators. The outputs of a computational collaborator or the collective may be appreciated by the human collaborator, even though they are not considered creative in a larger societal context. This relates to the concepts of personal and historical creativity, which distinguish historically remarkable creative artefacts from casual creativity [6]. Shneiderman [104] regards it is also important to provide support to small-scale creative activities. In practise some human collaborators acknowledge a clear contribution from the computational collaborator, for example, in creating new architectural designs [6], while others find the influence more subtle, such as influencing thought processes [22] or prompting towards new ideas (paper II).

From a societal perspective there appears to be some type of a bias against computational artists in general. Jordanous [69] considers the public critiques received by the Beyond the Fence musical, a musical advertised as the 'first computer-generated musical'. According to her, the critiques are in many cases focused on the involvement of humans in the creative process. This led her to believe there might be a bias affecting how computational participants' contributions to co-creative scenarios are recognised. Her empirical study found no significant differences in how outside evaluators evaluated the creativity of a computational system when it was considered as a computational collaborator, as part of a collective, or as a stand-alone creative system. Instead she found that outside evaluators were significantly more confident in evaluating the system as an individual, and some reported attributing the creativity of the collective system completely to the human. Although Jordanous suggests humans may be less confident in assessing the creativity of groups in general, her findings indicate that humans find it difficult to consider the creativity of computational collaborators or human-computer collectives. It may also be that the societal reception of human-computer co-creativity is somewhat contextual as the published opinions of the press appear more harsh than the privately shared opinions of the student participants of Jordanous' study. More studies on bias are needed to confirm her findings.

D'Inverno and McCormack [36] assess that the human collaborator's reception of a computational collaborator is linked to the societal reception of autonomous creative systems. In their view, humans seem unable to appreciate the artefacts produced by autonomous creative systems and the definition of creativity itself seems to change as autonomous systems achieve something new. They conclude that computational collaborators in art should not exist for their own ambitions.

Computational creativity designers should also recognise the potential negative effects of computational collaborators and the society at large. These include breaking social norms, or changing large cultural concepts such as education and employment [104]. The cost of new technologies may also lead to users being unable to afford participating in the newest creative trends [104]. Co-creativity may also be used to advance unethical goals, such as fake news campaigns [11]. Shneiderman [104] suggests participatory design practises and social impact statements to counter the potential negative effects of new creative technologies.

2.6 The Creative Product

The product refers to the outcome of the creative collaboration between the human and the computational agent. Traditionally the product perspective has focused on the type of the product and its evaluation [97]. In this thesis I also consider the contributions of the different collaborators to the product.

Products can be categorised by type. Classically this can mean the product's use, media of expression, utility and aesthetics [97]. In computational creativity the type of the product is closely connected to the domain of creativity in which the system works, and the medium of expression is affected by the choice of generative methods and output channels available for the system. In the case of the Poetry Machine, the co-creative product is the final poem produced from the original poem fragment composed by the Poetry Machine, through interactions with the human collaborator.

Evaluation of the creativity of the product is a classical theme in creativity research. Products are evaluated to recognise their creativity and separate them from innovations improving on existing ideas [97]. In computational creativity the product can be evaluated with different measures, such as quality, typicality, novelty (see e.g. [98, 99]), or surprise (see e.g. [49, 83]). Many of these evaluation measures can also be used during the generation of creative products. Surprise, for example has been used for directing both autonomous evolutionary search [52], as well as to pursue longer-term goals during co-creative processes [51].

Many evaluation measures are contextual and therefore relate to the creative press. In human-computer co-creativity, it becomes important to ask who must appreciate the qualities of the product in order for it to be considered good: is it the human or the computational collaborator, or some outside observer?

Using outside observers is common in computational creativity evaluation [67, 76]. In this thesis I consider the evaluation from the collaborators' perspective. In a fruitful collaboration between a human and a computational agent, the creative product will be appreciated by both collaborators. Yet in cases where the roles of the collaborators are asymmetric, we can not necessarily say that the end product will be appreciated by both. Issues related to achieving mutually appreciated results are discussed from a theoretical perspective in paper V. In general it may be more useful to ask the collaborator with a leading role, if they consider the outputs of the other to be of a high quality and have high utility for them. The human collaborator of the Poetry Machine is in such a leading position, and therefore we have considered the usefulness of the computational suggestions in paper II and quality of the end results in paper III, from the human collaborator's perspective.

Co-creativity can take place even when collaborators do not contribute equally. Contributions can differ either qualitatively or quantitatively. Contributions are qualitatively different if collaborators contribute to different aspects of the end product. For example, in human-human co-creativity, artists can participate in the same film production as a writer or an actor [1] or in digital art carry out the technical implementation of an artist's idea [17]. Qualitatively similar contribution is not a requirement for human-computer co-creativity either [128]. In human-computer co-creativity collaborators can participate in the same process, for example, via the generation or evaluation of artefacts.

Quantifying creative contribution is difficult, as it may not manifest in a visible way in the end product. For example, we calculated the percentage of words supplied by the Poetry Machine for Paper II, and found that two pupils used none of its suggestions in their final poems. However, the observation records showed that at times the children pointed to specific words, indicating they got an idea from them, even though they did not use that exact word in their final poem. Ultimately, Swartjes and Theune [110] consider that more important than visible contributions is that the creative collaborators accept each others' ideas as part of the space of possible creative products. This ties creative contributions to human experience. In Paper IV we examined creative contributions from this perspective, measuring such as the quality of other writers' ideas or subjective sense of ownership over the end product through self-reports filled in by the human collaborators. In our experiment, subjective ownership increased when the number of other writers and the quality of their ideas decreased, indicating that the amount and quality of others' contributions are linked to how human collaborators perceive their own contribution to a co-creative process. More information is still needed about the quality and effect of different contributions to the co-creative experience.

Chapter 3

Interaction Design for Human-Computer Co-Creativity

The purpose of interaction design is to help users to achieve their goals in a specific context [102, p. 317]. The creative context poses challenges to interaction design, as users behave in unorthodox ways [11, 105], and the requirements and measures for success are unclear [105].

Interaction design itself is also a creative activity [102, p. 317] and interaction design researchers attempt to influence actual product design in the wild by influencing engineers working on design tasks outside of academia [4]. This is achieved by gathering and disseminating design studies. A large body of design knowledge exists to support good design, illustrating what a good design process should be as well as practical tools for design.

In this chapter I describe how interaction design is conducted for human-computer co-creativity in practise. I start this chapter by considering the general activities related to the design process of human-computer co-creativity and proceed to illustrate current developments in design tools for human-computer co-creativity. In the final section I consider how this thesis has contributed to the design practise in human-computer co-creativity.

3.1 Design Process

Interaction design emphasises the involvement of users in the design process [102, p. 327] and follows an iterative design paradigm [20]. Many different interaction design processes exist, but most separate conceptual from concrete design phases [61]. The core activities in interaction design processes include establishing requirements, generating alternative designs, prototyping, and evaluating the results of the activities [102, p. 15].

We conducted an analysis of case studies on designing human-computer co-creative systems for Paper I and discovered a typical design process for human-computer co-creative systems. Our analysis was based on the ISO standard for human-centered design of interactive systems [62], which also describes a set of activities for typical interaction design processes.

Figure 3.1 compares the core activities of interaction design [102, p. 15] (left) and the activities listed in the ISO standard 9241-210 [62] (middle) with our resulting process (right). The figure shows that the core activity of establishing requirements is divided into two specific stages in the ISO standard. These include understanding the context of use and specifying user requirements. The activity is represented by one stage in the design process for human-computer co-creativity. Both the ISO standard and the design process for human-computer co-creativity combine the core activities of designing alternatives and prototyping into one step. These activities are followed by the evaluation activity in all frameworks. In the design process for human-computer co-creativity evaluation can be further divided into formative prototype testing and final evaluations. In this section I discuss how the specific interaction design activities and some general principles are reflected in human-computer co-creativity design literature.

3.1.1 Understanding the Context

Understanding the context of use allows interaction designers to ground their designs in the real needs of the users of their designs. False claims about users and their context need to be corrected at early stages of design when altering specifications is affordable and design solutions have not yet been limited [102, p. 37].

Important tasks for this activity include identifying all project stakeholders, their characteristics, goals and roles in the process, potential risks to them, and the physical and social environments in which the system operates [62]. Stakeholder participation ensures the appropriateness of the design, helps to manage expectations and fosters a sense of ownership of the finished design [102, pp. 322-323]. Balancing the needs of stakeholders in different contexts of use can be difficult [62].

The context of creative work is different from the traditional productivity-oriented context most interaction designers are familiar with [31, 105]. Traditionally the current context of use and similar systems have been used to understand the constraints under which the new system will operate [62], but the lack of examples makes this task difficult. It is also difficult to predict the effects of the collaboration on the stakeholders, as the societal role of human-computer co-creativity is not well established. Therefore it

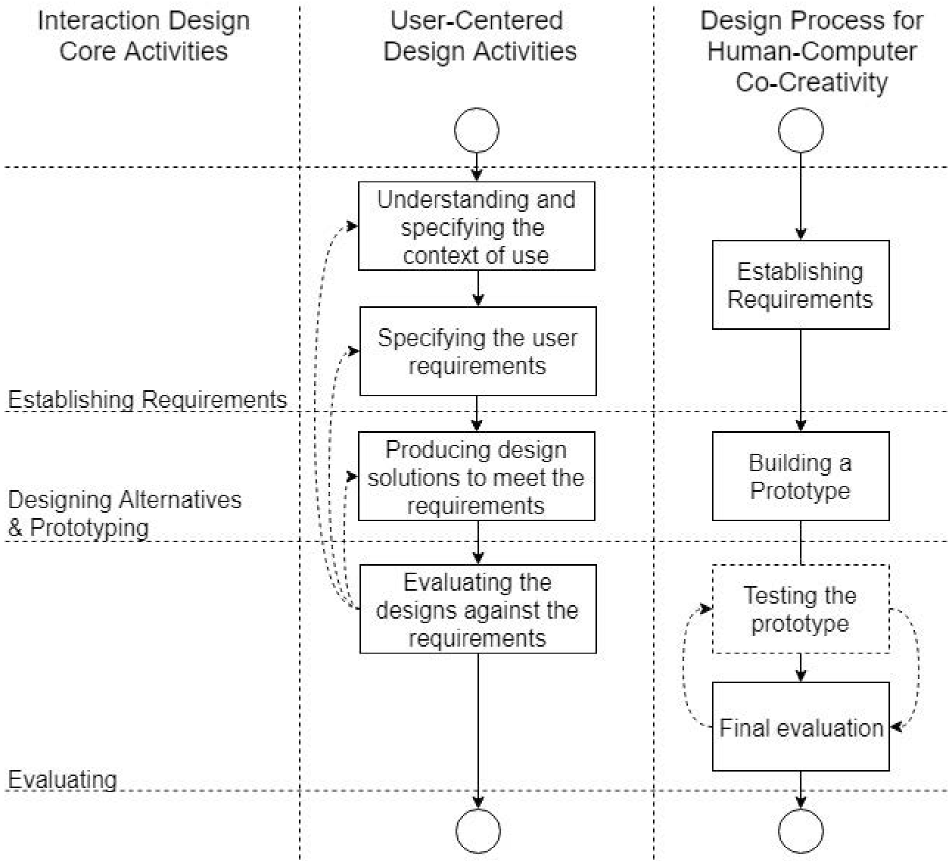

Figure 3.1: Interaction design core activities [102, p.15] (left), the ISO design activities [62] (middle), and the interaction design process for human-computer co-creativity according to the survey in Paper I (right).

is surprising how few details about the context are reported in the case studies reviewed for paper I. More studies are needed to establish good methods for understanding the context of human-computer co-creativity.

3.1.2 Establishing Requirements

Requirements describe how a design should perform and what it should do. Establishing a stable list of requirements for the design early on is important as fixing problems resulting from poorly established requirements is costly at later stages of the design process [102, pp. 353-355].

Establishing requirements can start with an initial set of requirements inherited from a previous product, or with no requirements at all [102,

p. 352]. Requirements should relate to the intended context of use and the objectives of the users [62]. Requirements should be supported by data, which can be gathered with various methods, depending on the nature of the task, the participants and available resources [102, p. 366]. Additional requirements can be derived from guidelines and standards, or organisational requirements and objectives [62]. Establishing requirements should be iterative: New data should influence old requirements [102, p. 353], and some requirements only emerge once an initial solution is proposed [62].

Solving trade-offs between contradictory requirements depends on the context and purpose of the design solution [62]. For example, user preferences may be less important in critical systems [62]. In human-computer co-creativity balancing strict requirements and softer preferences depends on the nature of the task: in autotelic, explorative creativity, hedonistic factors may be more important than in goal-oriented creativity.

3.1.3 Designing Alternatives

Designing alternatives involves conceptual design and physical design [102, p. 330]. Conceptual design provides a framework for general concepts and their interrelations for the design, including concepts users are exposed to via the design, and relationships between the conceptual design and user experience [102, p. 330]. Moving on to physical design, the designers need to decide what functions the design will perform, how the functions are related, what information needs to be available [102, p. 408] and what kind of interfaces best support interaction with the system [62].

At the conceptual stage, designers need to consider the division of work between the user and the system [62]. This step is important in human-computer co-creativity, as designers need to consider the role and impact of each collaborator in the human-computer co-creative process. Suggested roles can be found in Section 2.4.2.

Prototypes are an important tool in making design alternatives explicit so that they can be experimented with and improved upon [62]. Prototypes allow stakeholders to interact with design alternatives, assess their suitability [102, p. 390] and give feedback [62]. The accuracy of prototypes varies, from low fidelity prototypes made with cheap materials such as paper, to high fidelity prototypes that resemble finished products [102, pp. 391-395]. Prototypes can also vary in scope and detail [102, p. 398]. Starting with low fidelity prototypes is important, as too realistic prototypes may lead designers to lock in on a specific design or users to think that the prototype is going to be the final product [102, p. 398].

Typical difficulties in designing alternatives include the need to try out multiple solutions and to consider different parts of the system in an integrated manner [62]. In our experience (paper I) a major difficulty in human-computer co-creativity is in utilising low fidelity prototypes to facilitate discussions with users: It is difficult to construct useful low fidelity prototypes to communicate the capabilities and limitations of the computational collaborators to users at the early stages of design.

3.1.4 Evaluation

Evaluation should guide design throughout the design process instead of being just the final stage in it. Currently it appears that instead of continuous, iterative, evaluation throughout the design process, some case studies in human-computer co-creativity have approached evaluation as a sort of final assessment (see paper I). To improve the situation, more cost-effective methods useful for early prototyping and evaluation with human collaborators are needed. Interaction design evaluation for human-computer co-creativity is discussed in more detail in the next chapter.

3.1.5 Design Practise in Human-Computer Co-Creativity

The design processes is emergent making it difficult to manage [20]. To track the complex and non-straightforward process of design, Carroll [20] suggests establishing a design rationale, which describes the design solution or its design process focusing on the motivation of the design, its requirements, and how potential trade-offs were negotiated and specific features selected for the final design. Unfortunately, the documentation of design rationale in human-computer co-creativity design studies is limited. We found only a few documented design decisions in the case studies analysed for paper I. They suggest that when autonomous computational creativity methods are adapted into computational collaborators, the dependency of the computational collaborator from the human is increased in order to allow for meningful interactions. Further examples utilising more diverse interaction paradigms are needed to verify this finding.

The ISO standard [62] suggests six general principles for interaction design practise: it should be based on an explicit understanding of users, tasks, and environments, continuous user involvement, evaluation driven refinement, iterative process, addressing the whole user experience, and multidisciplinary development teams. It is interesting to consider how these factors have been realised in human-computer co-creativity design studies. In paper I, we noticed that although interaction design was common in the case studies, the design processes did not abide by these principles. For example the context of use, which relates to understanding the environment, was severely lacking in the investigated case studies and the nature of the processes was not iterative or iteration was limited. This is shown in Figure 3.1, where the dashed arrows describe typical iterations in the design process for human-computer co-creativity described in Paper I (right) and in the suggested iterations for the ISO design process (middle). Notably this limited iterative approach is similar to the recommended creativity support tool development process by Hewett et al. [56]. On a positive note, multidisciplinary project teams and user involvement seem to be fulfilled to some degree.

The ISO standard [62] also recommends planning the human-centred design process in advance to identify the relative importance of human factors for the project and potential risks from poor interaction design. Interaction design is clearly required in the area of human-computer co-creativity, but longitudinal studies are needed to assess its role and risks in the effectiveness of computational collaborators and their adoption. The ISO standard also calls for integrating interaction design as part of the product life cycle and the project plan for a new product. More information is needed on how human-computer co-creativity design can be integrated into computational creativity projects developing new methodologies.

3.2 Tools for Design

Interaction design is a holistic practise. Different design tools have evolved to offer designers guidance in different phases of the design process. Examples of design tools include design paradigms and frameworks, design principles, guidelines, and heuristics, design claims and patterns, and theories and models. They capture design knowledge distilled from design studies and offer theories in a form that is usable for design purposes.

Design tools vary in terms of scale [102, p. 55], specificity to a particular domain space [39, p. 259; 102, p. 55], and authority [39, p. 259]. In practise different tools may be in conflict and trade-offs between them are needed [39, p. 259].

A few different design tools have emerged to support the design of human-computer co-creativity. They offer another way to assess the status of design practise in the field. In this section I will discuss some recent tools contextualising them with tools for traditional interaction design and creativity support tool design. Theories and models for human-computer co-creativity are discussed in more detail in Chapter 5.

3.2.1 Design Paradigms and Frameworks

Design paradigms are shared assumptions, concepts, values and practices about design adopted by a community of researchers [102, p. 55]. Paradigms are rarely well defined [39, p. 186], while frameworks offer similar advice in a more structured format. Paradigms typically develop along technological advances, such as graphical user interfaces [102, p. 55] and the world wide web [39, p. 165].

Human-computer co-creativity has only recently started to emerge as its own field. Therefore the design paradigms used in the field are in a shift and not yet widely shared and documented. Currently a gradual shift away from creativity support tool paradigms is emerging, and new paradigms are established around different views on creativity.

The most important difference between the design paradigms for human-computer co-creativity and creativity support systems is in how collaborative creativity is viewed. Creativity support tool designers, such as Shneiderman [104, 105] often consider creativity as a social activity in their works, but for them collaborative creativity happens only between humans, and the purpose of interaction design is to facilitate communication between humans via well designed interfaces. In human-computer co-creativity collaborative creativity is seen as a process involving the computational system as a collaborator, and the purpose of interaction design is to facilitate interaction with it, not through it.

Design paradigms within the field of human-computer co-creativity can be differentiated according to how creativity is viewed by the designers: As a goal- and productivity-oriented activity, or as an autotelic, explorative activity. The idea that designing for creativity is different from designing for productivity is already present in creativity support system design literature [105] but autotelic, and explorative creativity have recently been highlighted in human-computer co-creativity literature (see e.g. [11, 13, 26]). Current research in human-computer co-creativity seems to exist in a spectrum between goal- and productivity-oriented creativity and autotelic creativity. How creativity is viewed on this spectrum affects its evaluation [72], and also rules for design: guidelines for a co-creative tool for interaction designers utilised by diPaola et al. [38] illustrate traditional usability goals, while design patterns suggested for autotelic and playful applications by Compton and Mateas [26] highlight different issues. Similarly Koch [74] suggested a framework for designing co-creative systems for design that is based on understanding the user's goals, while the patterns by Compton and Mateas focus on creative exploration of design spaces via offering a variety of options. These developments may be linked to the idea of Bray et al. [14] that computational creativity systems are gradually becoming available to non-expert users.

3.2.2 Interaction Styles

Interaction style paradigms describe how users interact with a product [102, p. 46]. Rogers et al. [102, p. 46] describe four general interaction styles for interaction design: in the *instructing style* a user issues instructions to a system. In the *conversing style* users have a question-reply type of interchange with the system. In the *manipulating style* users interact with the system by manipulating physical or virtual objects and in the *exploring style* users move through a virtual or physical environment. The styles are not mutually exclusive and help in formulating a conceptual model for the interface [102, p. 46]. I have used these styles to categorise different human-computer co-creativity interaction styles as presented in Table 3.1.

Three human-computer co-creativity interaction styles fit the instructing category: highly encapsulated systems, programmable interfaces [14] and operation-based interfaces [11]. In highly encapsulated systems the details of the computational collaborator's creative process are hidden and the human's participation is limited to parametrised abstract controls [14]. In programmable interfaces the users define their own set of commands with which to instruct the computational collaborator [14]. The operation-based interface offers similarly limited controls [11]. The design of operation-based interfaces considers typical interaction design issues, such as appropriateness of interface elements, visualisation, modularity, and offering multiple ways to edit and interact with the system [11].

The classic conversing style includes several suggestions for human-computer co-creativity. Request-based interaction is used to describe relatively open interfaces in which the computational collaborator appears autonomous and its role is more akin to a person [11]. Other suggested styles include an iterative interaction style focused on the refinement of ideas with a computational system and an additive style focused on the introduction of new ideas [22]. Young and Bown [131] have suggested several substyles for iterative interaction with improvisational agents: shadowing, mirroring, coupling and negotiation. All of these consider different degrees of collaborator participation and equality, starting from almost exact mimicry and culminating in an equal status. Design concerns for request-based interaction include how to exhibit consent, present relevant, clear, meaningful and diverse results, and how to handle ambiguity of requests [11]; these can also be considered for iterative and additive interactions.

	Classic Interaction Styles *Rogers et al.* [102, p. 46]			
	Instructing	Conversing	Manipulative	Exploring
Bray et al. [14]	Highly encapsulated systems		Direct manipulation	
	Programmable interfaces			
Bown and Brown [11]	Operation based	Request based		Ambient
Clark et al. [22]		Iterative		
		Additive		
Young and Bown [131]		Shadowing		
		Mirroring		
		Coupling		
		Negotiating		

Table 3.1: Interaction Styles for Human-Computer Co-Creativity.

Similar to manipulative style interfaces, the direct manipulation style of Bray et al. [14] offers a clear representation of objects the user can manipulate. Like exploring style interaction, the ambient interaction style suggested by Bown and Brown [11] considers interaction with embedded technologies. In human-computer co-creativity, ambient systems can be proactive and offer constant updates, assistant-like suggestions, or ambient information in the background [11]. Their design considers how to avoid workflow disruptions and how to be sensitive to the user's context [11].

Bown and Brown [11] argue that the semi-autonomous dynamism of co-creativity raises new interaction design issues. Thus interaction paradigms for human-computer co-creativity need to consider additional issues: Clark et al. [22] argue that in addition to the interaction structure, designers should also consider interaction initiation and intrusiveness. According to them, interaction can be initiated either by the human (pulling), the computer (pushing) or both. Interaction intrusion describes how well the computer's suggestions can be ignored, e.g. are they directly added to the creative product, or can they be deleted. Karimi et al. [72] consider interaction on a spectrum from computer-initiative to user-initiative domination, which affects the frequency of communication, and what is communicated. The concepts of interaction initiation, intrusion and domination offer interaction perspectives for augmenting the structure-focused view of traditional interaction design.

3.2.3 Design Rules for Human-Computer Co-Creativity

Design rules offer concrete guidance for design. In this section I consider design principles and guidelines. Design principles are general rules, focused on the psychological, computational and sociological aspects of design [39, p. 259], while guidelines describe more detailed rules for interface design [33; 39, p. 259]. Both are applicable to a wide range of designs, but guidelines have higher authority [39, p. 259]. Other interaction design rules, such as heuristics exist also for evaluating and designing systems [39, p. 282], but are not yet extensively adapted for human-computer co-creativity.

It seems that few general interaction design rules fit human-computer co-creativity directly. For example Norman's visibility principle seems to be at odds with designing systems that exhibit autonomous behaviour [13]. Therefore new models and interaction design techniques are needed for interaction with complex artificial intelligence [13]. A few rule sets for designing human-computer co-creative systems have emerged in recent years, including application specific design principles by diPaola et al. [38], general design principles by McCormack and d'Inverno [86], and general framework-type collections of design themes by Koch [74], and Bown and Brown [11].

These rules have mostly focused on interaction styles and specific interaction mechanisms [see 11 and 86] and they include only a few traditional interaction design themes [see 38]. Rules considering the design process itself are non-existent and there is scarcely any overlap between the themes discussed in the rules: Two of the rulesets (presented in [11] and [86]) discuss co-creativity beyond collaboration between one human and one computer as well as the need for playful interactions to facilitate creativity.

Some rulesets presented in creativity support tool literature appear also useful for human-computer co-creativity. These include rulesets for open-ended exploration [111], and some domain-specific guidelines. In the field of linguistic creativity Johnson and Carruthers [66] collected 14 requirements for a poem-writing tool, including general suggestions, such as allowing users to make quick notes, more poem-related suggestions, such as providing support for restructuring notes into verses and measuring syllables, and specific tools such as a thesaurus and a rhyming tool. Many of their suggestions seem appropriate also for co-creative poem writing. The Poetry Machine offers some of these, such as, separate writing spaces and a space for storing words or verses, poetic structure checking, themes and rhymes. However, since the purpose of the Poetry Machine is to be co-creative with the human, some supportive tasks, such as a thesaurus were left out and the Poetry Machine uses more finalised suggestions, such as concrete poem drafts to engage the user.

3.2.4 Design Patterns for Human-Computer Co-Creativity

Design patterns offer more specific design advice incorporating theoretical argumentation with illustrative examples [33, 109]. They attempt to combine a number of successful designs to offer invariant solutions to reoccurring problems in specific contexts [33]. They have multiple functions in a design process, including facilitating communication between stakeholders [39, p. 285], and use as educational tools, design structure suggestions, design rationale and organisational memory [33]. Patterns should be developed in a collaborative manner, but establishing them is difficult as there are no established criteria for recognising successful design [33; 39, p. 285].

Compton and Mateas [26] present design patterns for autotelic co-creative systems called "casual creators". To my knowledge, their attempt is the first to gather patterns for human-computer co-creativity. Therefore it is not yet supported by multiple examples or collaboration with other researchers. However, they still offer the most concrete design suggestions for human-computer co-creativity available. Their patterns promote easy, confident navigation of design spaces, focusing on finding objects that motivate the user and promote a sense of ownership. Compton and Mateas believe that casual users may be more willing to give control to a generative system than productivity-focused professionals. Thus their advice is suitable for researchers following an autotelic, exploratory creativity paradigm. Many of their patterns consider similar themes to those presented in creativity support systems guidelines. These themes include search and visualisation rules present in the work of Shneiderman [104, 105], simplicity and interoperability principles suggested by Resnick et al. [96], and themes for human collaboration and dissemination of ideas presented by both Shneiderman and Resnick et al. In a sense, the patterns of Compton and Mateas suggest that some principles already discovered in creativity support tool design are also applicable to human-computer co-creativity, although they might need different representations.

3.3 Contributions to Design Practise

The main contribution of this thesis to design practise is examining the design process for human-computer co-creativity as a whole. As far as I know, our paper I represents the only attempt at outlining a general design process for human-computer co-creative systems. It shows how human-computer co-creativity designers appear to mostly focus on concrete phases of design and do not sufficiently emphasise early design.

In our investigation of design processes for Paper I, we also found that in many case studies, design rationale and trade-offs made during design are poorly documented, making it difficult to conduct reviews on design practise and design processes in human-computer co-creativity. It also appears there is a lack of methods to support early, conceptual design. We also found that constructing human-computer co-creativity systems on the basis of computational creativity methods may require a re-design of the methods to accommodate for the user's role within the co-creative process. These findings are considered in more detail below.

As the work outlined in this thesis began, much less literature on how to design human-computer co-creative systems was available. Due to the lack of literature outlining the design process for human-computer co-creativity we followed the general design principles suggested by the ISO standard for user-centered design in our own design process. During the process we found it difficult to apply the principles of early testing with users to our work, as we could not think of a way to use low fidelity prototypes to communicate the capabilities and limitations of the computational creativity methods powering the Poetry Machine to our young users. This may be a context-dependent finding, but as early design is seldom reported, I consider there is a need to further support early phases of design in order to enable early evaluation and iterative testing.

I have considered the human and the computer as collaborators throughout my thesis work. This appears already in our paper I, in which we used Wiggins' Creative Systems Framework [124, 125] to illustrate the capacities of the human and the computational collaborator. This approach allowed us to notice how existing computational creativity methods may need to undergo changes in order to better fit the increased need for interactivity in co-creative applications. Although the work conducted for this thesis does not include guidelines, patterns, or other forms of design rules, we extended the Creative Systems Framework into a model that can be used also as a design tool for human-computer co-creativity. The model is described further in Paper V and Chapter 5.

Chapter 4

Evaluating Human-Computer Co-Creativity

The success of human-computer co-creativity depends on contextual factors, including the characteristics of the collaborators and their creative task and environment. Interaction design evaluation takes into account the ambiguities related to evaluating the success of complex tools in complex contexts, thus offering a good methodological basis for evaluating human-computer co-creativity [8]. Interaction design evaluation methods can also be used to evaluate the effects of the collaboration, and assess the roles of the collaborators within the context.

Two evaluations of the Poetry Machine were completed during this thesis project. Their procedures and results are described in papers II, III, and IV. The first paper describes the evaluation of the first interactive prototype of the Poetry Machine, while the two latter ones describe an evaluation of the final prototype of the Poetry Machine. They describe two interaction design-based approaches to human-computer co-creativity evaluation: usability and user experience evaluation.

This chapter considers general issues related to human-computer co-creativity evaluation, illustrating a background for the evaluations conducted for this thesis. I follow a structure based on Aristotle's six questions, why, what, who, when, where and how, known as the *septem circumstantiae* [107]. Partial sets of these questions have been previously used to plan evaluations in interaction design [102] and evaluations of the creativity of human-computer co-creativity [72].

4.1 Why to Evaluate

The rationale for evaluation depends on the primary goals of the human-computer co-creativity design project. There are three goals for computational creativity projects [76] that can also be applied to human-computer co-creativity: artistic goals, including the improvement of personal art practise, design or engineering of new applications, and scientific experiments intended to advance the field. Evaluating artistic goals is outside the scope of this thesis. Rationale for design evaluation and scientific evaluation are considered underneath.

Evaluation is a core element in designing human-computer co-creativity applications. It ensures that the design process progresses towards a successful end result and the final design meets the requirements set for it. Offering feedback for the gradual improvement of design is one of the core purposes of both interaction design [20] and computational creativity evaluation [67]. In both fields, evaluation helps in discovering the strengths and weaknesses of different systems [56, 67], which other researchers can learn from [67]. Interaction design-based evaluation can also be used for collecting information on user needs and establishing baselines [62].

Sound, empirically grounded evaluations are also the basis of science. They connect theoretical claims into scientifically measurable events and transform designs to incremental scientific progress [8]. In computational creativity, evaluation is a way to show that specific criteria for creativity have been met [67, 76] and to demonstrate and track the progress of a system [67]. Computational creativity and interaction design methods can also be combined to answer specific, scientific questions about human-computer co-creative systems, such as how to distinguish them from creativity support systems [72] and what are the roles of the collaborators [130].

The goals of the evaluations conducted for this thesis have been to gather feedback for improving the Poetry Machine and to improve the scientific practise of evaluating human-computer co-creativity by experimenting with different evaluation methods and metrics.

4.2 What to Evaluate

Evaluation can be used to answer different questions about human-computer co-creativity. These questions are typically related to the creativity in and interactions with a human-computer co-creative system.

Thorough evaluations of creativity should consider all of the four perspectives, *person*, *process*, *product*, and *press*, presented in Section 2.2

[67, 68, 76]. The *person* perspective in human-computer co-creativity has considered the creativity of the human collaborator [72], and how to attribute creative agency within a co-creative system (see e.g. [9, 82]). In computational creativity, the *process* perspective is considered useful for providing guidance to designers, and the *product* perspective is relevant for systems intended to produce useful results for humans [76]. These two perspectives have also been named as potential targets for human-computer co-creativity evaluation [72, 128]. The *press* perspective sets the context in which other evaluations are conducted [76]. So far only Jordanous [69] has worked on this perspective of human-computer co-creativity. She approached the *press* perspective through assessing the confidence and bias of humans evaluating human-computer co-creativity.

The evaluation of interactions with a co-creative system mainly considers *user experience* and *usability* [8], but interaction evaluation could also be used to assess the impact of the computational collaborators [128], or to address specific scientific questions. *User experience* measures a person's perceptions and responses related to the use or the anticipated use of the design, including all of the user's emotions, beliefs and preferences [62]. *Usability* is a major factor of user experience [75] focused on the extent to which a design can be used by specific users to achieve specific goals in specific contexts with effectiveness, efficiency and satisfaction [62].

Questions related to creativity and interaction can also be combined when evaluating human-computer co-creativity [8]. Elements affecting user experience could be considered through the four perspectives suggested for computational creativity. In addition to evaluating the effects of the perceived qualities of the computational collaborator, the co-creative process and the quality of the product, user experience provides an interesting view to the *press* perspective through the evaluation of the anticipated co-creative process.

The evaluations performed in papers II, III and IV focus on evaluating the usability of and user experiences with the Poetry Machine. In the user experience evaluations we have also considered subjective creativity and self-expression as potential factors.

4.3 When to Evaluate

Evaluating human-computer co-creativity can be divided into *formative* and *summative* evaluation depending on the purpose and timing of the evaluation. *Formative* evaluation focuses on gathering constructive feedback that can be used to improve computational creativity methods [67], or interaction design prototypes [102, pp. 437] during the research project. In design projects, early formative evaluation is important to avoid costly mistakes [62]. *Summative* evaluations provide summary judgements of the creativity of computational creativity systems [67], and can be used to assess the success of a finished design [102, pp. 437]. Summative evaluation can also be used when assessing the scientific goals of a project.

Summative and formative evaluation can also be conducted during the co-creative process and after it in order to improve it: Karimi et al. [72] suggest using *formative feedback* to describe the feedback given by the collaborators to each other during a creative process and *summative feedback* as the feedback given after the process to improve the process in the future.

The evaluations in paper II describe formative evaluations of the first Poetry Machine prototype, while the evaluations in papers III and IV summatively compare co-creative user experience with the Poetry Machine to working with a human or with a human and the Poetry Machine.

4.4 Who Should Evaluate

Human-computer co-creativity can be evaluated either by the participants of the co-creative process, or by evaluators not participating in the process, including the system's designer, domain experts, or laymen. Who should evaluate depends on what is being evaluated and at what stage of the design process the evaluation takes place. Different evaluators can also contribute differently to the evaluation process. For example, in mixed-initiative co-creativity for games, the contributions of evaluators can be divided into direct and indirect [129, p. 198]: Direct evaluators, such as designers have the power to make decisions based on the evaluations, while indirect evaluators, such as human players or AI agents, provide feedback to the desicion makers via playtesting and subjective reports.

The main evaluator of the co-creative system is typically the human collaborator following a procedure outlined by the researchers (see e.g. [8, 72, 130]). The evaluations in this thesis were all conducted with human collaborators of the Poetry Machine. Human collaborators typically evaluate interaction, but they can also evaluate aspects of the computational collaborator (see e.g. [122]), the co-creative process [72], and its product (see e.g. [22, 46]). They can participate in both summative and formative evaluation, or evaluate their anticipated user experience prior to engaging with a co-creative system. Human collaborators provide a way to study real interactions with a system, but such studies can be expensive [102, pp. 441] and recruiting a representative group of users can be difficult.

Karimi et al. [72] hypothesise that computational collaborators could also participate in formative evaluation through self evaluations, or in negotiation with human collaborators. Koch [74] suggests they could also evaluate their human collaborator in order to adapt to their needs during the creative process. Currently the participation of computational collaborators in evaluation is limited by technology.

In computational creativity the complete evaluation is often conducted by the designers of the system due to limited resources [67]. To alleviate concerns of partiality, evaluation should be transparent [67]. In interaction design, this also considers documenting how evaluation data is used to guide design decisions.

Experts and laymen can also participate in human-computer co-creativity evaluation, judging for example the end product. However, use of human judges in evaluating creativity is challenging due to difficulties in explaining opinions, the large number of evaluators required for robust evaluations [67], possible bias against computational creativity [67, 76] and susceptibility to contextual and superficial factors [76]. Also, experts and laymen are not equally equipped to participate, as expertise and familiarity with similar systems affects evaluation results [67] and evaluator confidence [69]. Experts can also be used as a cheap and quick way to identify usability problems and predict user behaviours with design prototypes [102, pp. 441]. In human-computer co-creativity this approach is, however, currently limited by the lack of suitable evaluation heuristics.

4.5 Where to Evaluate

Interaction design methods cover a range of different settings including strict laboratory methods and natural, ‘in the wild’, methods [102, pp. 436]. Studying interaction phenomena in the laboratory offers more control over the use situation and reduces potential outside influences and distractions [102, pp. 437-438], whereas studies conducted ‘in the wild’ demonstrate how people use different technologies in their intended setting and help capture the real context of use [102, pp. 441].

‘In the wild’ studies are often expensive and difficult to conduct as anticipating where and when interesting phenomena will happen is difficult [102, pp. 440-441]. Therefore controlled laboratory studies may appear more appealing to researchers, especially when the goals of the project are scientific. However, creativity is difficult to study in laboratory conditions, which may reduce spontaneity [42] and fail to capture temporal aspects of creativity [3]. Collaboration is also difficult to study in a laboratory [56].

Thus I recommend a hybrid environment combining parts of the laboratory and ‘in the wild’ approaches to evaluate human-computer co-creativity.

The evaluations in this thesis were conducted at school, where we build a sort of hybrid evaluation environment: We brought our own equipment, outlined the task and carefully recorded our participants, while they enjoyed the peer support and familiarity of the real context. Details of the evaluation environment are discussed in paper III.

4.6 How to Evaluate

To conduct an evaluation, researchers need to select methods and metrics that fit the goals and context of the study. It is impossible to provide a complete review of interaction design and computational creativity evaluation methods, metrics and their applicability to human-computer co-creativity in this thesis introduction. Instead I focus on general issues related to the selection of methods and metrics and lessons learned when applying specific methods in evaluating human-computer co-creativity.

4.6.1 Selecting Methods

Combining different evaluation methods is recommended to gain a rich understanding of the usability and user experience of a system [102, pp. 442] and to alleviate method weaknesses [56]. In human-computer co-creativity evaluation, methods can also be combined across different fields [72]. To select a suitable combination, methods need to be compared in a meaningful way.

The choice of methods can be narrowed down by considering the type of data needed to answer the research questions. Qualitative data provides useful feedback for formative evaluation [67] and can be used to understand why users behave in certain ways [56]. Quantitative data is good for comparative, summative evaluation of different systems, or to show gradual progress between prototypes. However, quantitative evaluation of creativity [76] and aesthetics [11] have been criticised. Combining quantitative and qualitative data allows for grounding quantitative evaluation criteria in practise [67], capturing significant differences between systems, and explaining the effects [56].

Methods can be further compared across different fields by three criteria: scientific quality, usefulness and practicability (rows in Table 4.1). These criteria are based on meta-evaluation criteria for computational creativity [70], usability [54], and user experience [118] (columns in Table 4.1). The criteria and their background are defined in full below.

	Computational Creativity Jordanous [70]	**Usability** Hartson et al. [54]		**User Experience** Vermeeren et al. [118]
Scientific quality	Correctness	Thoroughness	Effectiveness	Scoping
	Faithfulness (as a model of creativity)	Validity		Scientific quality
		Reliability		
Usefulness	Usefulness	Downstream utility		Utility
	Generality			Specificity
Practicability	Usability	Usability		Practicability
		Cost effectiveness		

Table 4.1: Criteria for evaluating Human-Computer Co-Creativity evaluation methods (rows) based on domain-specific criteria (columns).

Scientific quality describes the scientific robustness of the method including the validity of results, thoroughness of the investigation and the method's independence from contextual factors. In computational creativity evaluation, scientific quality considers the accuracy and comprehensiveness of the evaluation (correctness), and how faithfully the evaluation captures the creativity of a system [70]. In usability evaluation, it considers the method's independence from evaluators (reliability), and statistics calculated for the coverage (thoroughness), proportion of false positives (validity), and their comprised effectiveness [54]. The user experience meta-evaluation criteria also consider the validity and reliability of a method. Additionally scoping considers how well the method is able to elicit information across different dimensions of user experience, e.g. emotion [118]. In addition to the criteria shown in Table 4.1, interaction designers also consider the possible biases a method has and the generalisability of the results [102, pp. 471-472] when considering scientific quality. The quality of our user experience evaluation results is considered in paper IV.

Usefulness refers to the usefulness of the data produced by the method and its applicability to different contexts and different parts of the design process. The usefulness of the results for different stakeholders is considered by both, computational creativity [70] and user experience [118] meta-evaluation criteria (usefulness and utility). So is the generalisability of the method across different domains (generality) [70] or different groups of users (specificity) [118]. The usability meta-evaluation criteria consider the usefulness of the method in different stages of the design process (downstream utility) [54]. The usefulness of our usability evaluation results were considered in paper II.

Practicability describes the ease of use and usability of the method with respect to time needed for both learning the method and conducting successful evaluations with it. The general ease of use of the methodology is considered in the meta-evaluation criteria across all domains [54, 70, 118]. Usability and user experience meta-evaluation criteria add to it considerations related to the method's cost effectiveness, requirements and expertise required, and considerations about users' motivation to participate in evaluations conducted with it [54, 118]. The practicability of our usability evaluation methodology was considered in an additional publication [71].

4.6.2 Deriving Metrics

Like computational creativity in general [67], human computer co-creativity suffers from the lack of clear evaluation criteria. Some studies (e.g. [63]) have used general usability evaluation questionnaires such as the SUS [15] or a creativity support tool-specific evaluation questionnaire called the Creativity Support Index [21], to evaluate human-computer co-creativity quantitatively. However, specific criteria for evaluating human-computer co-creativity are scarce.

Therefore many researchers use ad-hoc criteria to evaluate human-computer co-creativity. This approach has been criticised in computational creativity for the lack of justification for the criteria and their unknown relationships hindering the evaluation of the dimensionality of creativity [76]. Yet, even computational creativity is domain-specific and researchers have to carefully consider what best describes creativity in their domain [67].

In addition to being well defined and having a basis in literature, a good metric also has to be operationalised in a way that allows measuring it in practise. Bown [8] discusses traditional computational creativity metrics including value and novelty and notes that in human-computer co-creativity they become highly subjective. This promotes an approach looking at different metrics as subjective, user experience-based measurements.

In paper III we consider how criteria from evaluating computational creativity, creativity support tools, and user experience can be combined for comparing the experiences of human collaborators participating in different co-creative processes. Paper IV presents our results, including a preliminary evaluation of the correlations between the different metrics. Together the two papers describe early work on introducing new criteria for measuring human-computer co-creative experiences.

4.6.3 Practical Issues in Evaluation

The work conducted in this thesis has focused on documenting the practises surrounding human-computer co-creativity evaluation. These practises have rarely been discussed in human-computer co-creativity evaluation case studies, which mostly focus on evaluation results. In Paper II, we apply an evaluation planning framework called DECIDE by Rogers et al. [102, pp. 455] on evaluating the usability of the Poetry Machine. In Paper III we explain how user experience can be evaluated within the human-computer co-creative process. Throughout the evaluation studies presented in Papers II, III and IV, we have also reported practical issues in the use of specific interaction design evaluation methods in human-computer co-creative contexts as illustrated below.

We found that selecting suitable evaluation tasks for the creative context is difficult: in our first evaluation, described in Paper II, we attempted to promote creative thinking by having a very general evaluation task; "write a poem". This proved difficult for the young children who participated in the experiment alone and they needed some guidance, such as suggesting a topic. Then again, in our later, comparative evaluations described in Papers III and IV, we found that restricting users to a more specific predefined task seemed to limit their creativity. Users did, however, benefit from mentioning a sample goal, such as writing a poem to congratulate a friend who likes animals. It would therefore appear that suggesting specific topics helps to explain a creative task and leverage creativity, but enforcing the task, especially if the task is repeated during the evaluation, limits creativity.

We also found that it was difficult to analyse our observations of the children who participated in our usability evaluations, as facial gestures typically classified as negative seemed to represent concentrated immersion instead of usability problems. This may be an effect of the creative context, where other traditional measures used to capture usability problems, such as time, have been found to indicate positive immersion instead [19].

Our experience evaluations overall also demonstrated the need for using multiple methods to evaluate user experience in co-creative contexts: In Paper III we examined how interviews and observation supported quantitative data collected with questionnaires. We found that our 10-11-year-old users seemed in general to understand the questions and be able to answer them truthfully. Thus it appears that the differences between observed enthusiasm and questionnaire results in part demonstrate that it is very difficult to capture the creative experience by observation only. Subjective, internal reports of the experience are needed to be able to consider the user's co-creative experience in full.

4.7 Contributions to Evaluation

Two evaluations were conducted on the Poetry Machine in the course of this thesis. These evaluations considered different factors related to human-computer co-creativity, focused on the planning and conducting of evaluations in practise as well as suggested a number of evaluation metrics for human-computer co-creativity.

Paper II describes how a usability evaluation for the Poetry Machine was planned and conducted in practise. Usability is a common evaluation target in human-computer co-creativity. Our evaluation focused on formative evaluation. Multiple methods were combined in the evaluation. These methods offered data on both the practical usability problems of the prototype as well as the children's feedback on more conceptual issues, such as how to improve the quality of the suggestions generated by the Poetry Machine. The evaluation results were useful for improving the Poetry Machine. Their results were quite thorough, as indicated in [71].

In Paper III we compared three co-creative writing experiences: writing with the Poetry Machine, writing with a friend, and writing with the Poetry Machine and the friend. The paper describes ten quantitative evaluation metrics, related to seven evaluation themes that were derived from computational creativity evaluation, creativity support systems and interaction design. The paper offers a view on how evaluation metrics can be justified with arguments from literature, and how they can be grounded by using qualitative data. In Paper IV, we consider the full results of our study and correlations between the metrics. We found that a comparative method, ranking different co-creative experiences along the selected criteria, elicited many statistically significant results. Paper IV also shows how to consider the bias and reliability of our results. We found that ownership is a promising new metric for evaluating human-computer co-creativity, while creative self-expression and ease of finishing writing correlate with all other metrics.

Our evaluations also discovered that the task used in evaluating human-computer co-creativity is best expressed in the form of a specific example goal. However, restricting participants to this goal may limit their creativity. The context of the evaluation was discussed extensively in our papers, with an emphasis on factors with potential effects on creativity, such as a relaxed atmosphere. We also found that it is difficult to rely on observational data when evaluating human-computer co-creativity and that self-reports are important for final conclusions.

Chapter 5

Modelling Human-Computer Co-Creative Processes

Models provide simple concepts and language for discussing complex phenomena. In creativity research, models are constructed to identify, develop and facilitate creative thought [88]. In computational creativity, models support the comparison [124] and evaluation of different systems and in interaction design they offer tools for design by connecting different theories [109] and offering them in a simplified form [102, p. 55].

The modelling approach taken in this thesis focuses on the human-computer co-creative process and approaches it from a computational perspective. Our model, described in paper V, is based on Wiggins' [124, 125] Creative Systems Framework, which describes computational creativity as a search in a universe of possible creative concepts. Wiggins' framework in turn is based on Boden's (e.g. [6]) concepts of exploratory and transformational creativity, which consider the exploration of structured conceptual spaces in order to generate new ideas and the transformation of these spaces to allow for the generation of new types of ideas. Our model formalises human-computer co-creativity through the concepts of exploratory and transformative creativity and operationalises them as an iterative creative search between a human and a computational agent. The model offers a tool for analysing different design options by illustrating some potential issues arising during human-computer co-creative processes and describing potential roles for the collaborators.

I will start this chapter by examining different approaches to modelling human creative processes, computational creativity and human-computer co-creativity, focusing on the main ideas behind them. I will then examine the main themes of our model. Finally I will summarise the benefits and weaknesses of our approach and how it complements the other approaches.

5.1 Modelling Human Creativity

There is no consensus over modelling human creativity: some authors propose that the creative process is entirely individual [42], while others suggest there are also some universal components to it [88]. Lubart [81] concludes that there are both general and domain-specific models and in some cases general models can be adapted into domain-specific models by replacing some subprocesses. Several different styles for modelling creativity exist, such as cyclic processes [60, 81], modelling distinct sets of actions taken in the creation of the artefact [81], and cognitive process models [88]. The most popular style of model is stage-based and linear [60].

Several different stage-based models exist, suggesting different steps for the creative process. Common steps include problem finding, information finding, idea finding and solution finding, which have been used, for example, to compare the support offered by creativity support tools [120]. Stage-based models have also been used to describe human-human co-creative processes. Suggested steps for human-human co-creativity include idea generation, development, finalization, with closure, and evaluation [91], or concept creation, construction, and evaluation [17]. Stage-based models may also be useful for modelling human-computer co-creativity, since they seem to be easy for users to understand as indicated by a recent study [22] in which users suggested adding idea generation and evaluation as specific steps to the human-computer co-creative process in order to improve it.

Currently, models of human creativity and creative processes are being under-utilised in design: a recent study of creativity support tools found that designers typically only consult a limited number of theories [120]. To aid the design of creativity support tools, Shneiderman [104, 105] attempted to combine several theories of human creative process into three perspectives. His perspectives, structuralists, inspirationalists, and situationalist, emphasise different approaches to creativity, with the first highlighting a stage-based approach, the second an intuitive approach and the final a social approach. To help designers, Shneiderman lists specific support mechanisms for users fitting each specific perspective. He recommends combining different perspectives in order to build more useful creativity support tools.

Shneiderman [105] also notes that creativity support systems already enable completely new creative processes. Therefore theories about human creativity alone are not sufficient to describe the joint creative process with a computational collaborator.

5.2 Modelling Computational Creativity

Models for computational creativity include descriptive models for the evaluation of specific properties of the computational creativity process and models for the process itself. According to Lamb et al. [76], the most popular theory for the computational creative process is Boden's theory of conceptual spaces. Wiggins has operationalised this theory as a search in his Creative Systems Framework [124, 125]. The Creative Systems Framework itself has been further adapted to operationalise specific traits of creativity, such as curiosity [50], or model other theories of creativity, such as divergent and convergent thinking [115].

Boden [6] defines three types of creativity: combinational, exploratory, and transformational. The first type generates improbable combinations of familiar ideas. The second constructs novel ideas through the exploration of structured conceptual spaces. The final type transforms the dimensions of this space in order to create new ideas that could not have existed before.

In Wiggins' [124, 125] model, Boden's exploratory creativity is described as a search in a universe of possible creative products $\mathscr{U}$. The space is traversed through a traversal function defined by ruleset $\mathscr{T}$. The eligibility of found candidate products is evaluated with respect to two rulesets $\mathscr{R}$ and $\mathscr{E}$, which describe the validity and quality of the the candidate respectively. These rulesets define what the computational agent considers as a valid example of a product in a specific domain, and what it considers as valued (or aesthetic). The computer is only able to generate a product if the traversal rules can reach concepts, which it considers to be both valid and aesthetic within the universe. Wiggins [124] describes a number of different situations in which suitable concepts can not be found because of the properties of the rulesets. He argues that the mismatch between $\mathscr{R}$ and $\mathscr{E}$ is the driving force behind transformational creativity, which occurs as the system attempts to reconcile this mismatch by changing $\mathscr{R}$, $\mathscr{E}$, or $\mathscr{T}$.

The Creative Systems Framework itself gives some concepts for analysing computational creativity systems. We apply it to the analysis of human-computer co-creative systems in Paper I. It has also been used to guide the design of autonomous computational creativity systems (see e.g. [48]). However, for building autonomous computational creativity agents, Ventura's [117] recent algorithmic model for building computational creativity systems may provide more useful suggestions. Neither model, however, considers interaction with a human.

5.3 Modelling Human-Computer Co-Creativity

Few different models have been suggested for human-computer co-creativity specifically. These include domain-specific models, such as an iterative stage-based model for story generation [110] as well as more general models including an interaction design-oriented model called a Four-Strategy Model of Creative Interaction [115] and a cognitive model called an Enactive Model of Creativity for Collaboration and Co-Creation [30]. The Four-Strategy Model of Creative Interaction [115] describes different interaction strategies based on combinations of implicit and explicit thinking with convergent and divergent creative processes. The Enactive Model of Creativity for Collaboration and Co-Creation [30] describes co-creativity as a participatory sense-making process. In this thesis introduction I focus on the enactive model [30], as it is intended to be domain-independent and focuses on the creative process.

The *Enactive Model of Creativity for Collaboration and Co-Creation* [30] is a model based on the enactive paradigm of cognitive science. According to the model's creators the enactive paradigm considers cognition as a dynamic, sensorimotor loop, where actions and perceptions are inseparable. The model focuses on the first-person experience and the creative agent's awareness over its own intentions. It describes creativity on a low level, avoiding high-level cognitive mechanisms, such as verbal interaction with the human collaborator. The authors argue that the model includes the crucial environmental feedback loop, which is left out in transformational creativity-based models.

Three factors contribute to the improvisational emergent creativity described by the enactive model [30]: the intentional state of the agent, its skills and capacities, and the perceived affordances of the environment. The agent is able to direct its attention between the properties of the artwork and its mental model and choose suitable actions through perceptual logic. The design advice given by the model focuses on how to encode different levels of perceptual logic to describe the agent's awareness of local, regional, and global aspects of the artwork, inputs of the user, and its own potential activities.

The enactive model [30] seems capable of explaining a number of interesting phenomena, like learning and the effects experience and distraction have on creativity. It has also been used to deliver interesting computational collaborators, such as the Drawing Apprentice [31]. One of the most important decisions left to the designer using the enactive model is to decide how to encode the different layers of perceptual logic needed to guide the actions of the system. The creators of the model present examples from three domains: design, visual arts and music, but it is unclear how suitable the model is for domains with less clear rules of perception, such as language.

5.4 Human-Computer Co-Creativity as a Search

Our extension of Wiggins' Creative Systems Framework in Paper V describes human-computer co-creativity as an iterative process in which the human and the computational participant take turns in modifying a joint creative product. The details of the model, including its mathematical formulation, can be found in Paper V. In this thesis introduction I focus on describing the properties of the model and the terminology it provides for analysing and designing human-computer co-creative systems.

The model defines two modes for creative collaboration: alternating-co-creativity and task-divided co-creativity. *Task-divided co-creativity* allows the human and the computational collaborator to assume distinct duties, such as concept generation, or evaluation in the course of the collaboration. It can happen between *incomplete agents*, which individually do not have the full capacity to define, traverse and evaluate creative spaces. Task-divided co-creativity allows for formalising three different roles for creative collaborators: concept definer, concept generator, and concept evaluator.

Alternating co-creativity assumes that the human and the computational collaborator operate in the same creative universe, but each has their own sets of rules for traversing the space, and evaluating the validity and quality of concepts in it. The agents are *complete* in the sense that they could also operate individually. Alternating co-creativity can be *symmetric* or *asymmetric* depending on whether or not the collaborators are allowed to skip turns in the co-creation.

Several different issues arise when collaborators try to assume symmetric, alternating co-creativity: First of all, to be able to work on the joint task, the collaborators need to operate in the same universe, otherwise a *universal mismatch* will occur. In order to be able to discuss similar concepts, the collaborators need to be able to agree what constitutes a valid artefact in the process, or a *conceptual mismatch* will occur. Differences in rulesets for evaluating the aesthetics of the products can also lead into *artistic disagreements*. Finally, if one of the collaborators is unable to continue the creative search from the product it receives from the other, it suffers from *generative impotence*. When designing computational collaborators, designers need to consider each of these issues and how to address them. Some of the issues may be best solved through transformative action, allowing the collaborators to extend or limit their search to more fruitful conceptual spaces. One way to consider these transformations is to reflect on whether the computational collaborator should assume a *pleasing* or a *provoking* role with regard to the suggestions proposed by the human collaborator.

5.5 Contributions of Human-Computer Co-Creativity as a Search Model

Our model describes co-creativity as search, using concepts familiar to computational creativity researchers. The co-creative process is described as a continuous, iterative operation, which does not fit any typical human creativity processes I am aware of. However, task-divided co-creation could also be formulated as a stage-based iterative model with generative and evaluative stages.

The concepts of universal and conceptual mismatch, artistic disagreement and generative impotence can be used to identify key points in interaction, which need to be addressed in the design of co-creative systems. The terminology introduced also allows the comparison of different kinds of computational collaborators and assessing their strengths and weaknesses.

The approach of our model is different from the enactive model of creativity [30], which does not suggest specific problems for co-creative processes. Considering the enactive model in terms of our model, the actions of the agent would roughly correspond to the traversal rules in our model and the perceptual logic to the evaluation rulesets $\mathscr{R}$ and $\mathscr{E}$. The issues of conceptual mismatch, artistic disagreement and generative impotence would then correspond to situations in which a perception-action pairing would be missing, or the perceptual logic would result in poor quality results (in the human collaborator's opinion). Our model could help conceptualise and address these issues during the design of the system.

The two models could be combined to reach a more intentional and autonomous computational collaborator; the enactive model already suggests how the computational collaborator can pursue its own artistic goals in the form of intentions. Our model could be helpful in achieving flexible sets of perceptual logic, which could be adapted following our analysis of artistic disagreements and conceptual mismatches. However, there is no meta-level information sharing or communication in either model beyond the shared artefact. To increase the intentionality of the system and to enable goal-oriented co-creative behaviour, adding a model for sharing meta-level information, such as evaluations, would be critical.

Chapter 6

Conclusions

In this thesis I investigate human-computer co-creativity from three perspectives: design, evaluation and modelling. In addition to a theoretical perspective, I assume a practical, interaction design perspective and consider these themes through case studies conducted with a co-creative poetry writing system, called the Poetry Machine.

I define human-computer co-creativity as creative collaboration with at least one human and one computational collaborator. In this thesis introduction, I have examined the characteristics of the creative collaborators involved in the co-creative process, the reception of human-computer co-creativity and its context as well as the characteristics of the products created. These concepts, adapted from Rhodes [97], give a common background for discussing the design, evaluation and modelling of human-computer co-creativity.

This thesis approaches the design of human-computer co-creativity systems through a case study describing the design process of the Poetry Machine system. This design process and other case studies are used to elicit a general process for designing human-computer co-creative systems. Chapter 3 summarises our findings on interaction design practise in human-computer co-creativity based on this general process. Papers II, III, and IV outline two case studies on the evaluation of the usability and user experience of the Poetry Machine system. Chapter 4 provides the background for comparing our evaluations to other case studies in the field. Modelling of human-computer co-creativity is described from a theoretical perspective in Paper V. It presents our iterative model of human-computer co-creativity as search, which I compare to different approaches to modelling creativity in Chapter 5.

6.1 Answers to Research Questions

I consider four research questions in this thesis: how does the design process of human-computer co-creative systems differ from typical design processes of interactive systems? How can qualitative evaluation guide the design of a co-creative system? How can quantitative evaluation be used to compare different co-creative processes in a meaningful way? And how can the human-computer co-creative process be described in a way that can be used to guide design decisions?

My first research question, *how does the design process of human-computer co-creative systems differ from typical design processes of interactive systems,* is addressed in Paper I and Chapter 3. It appears that interaction design methodology is well utilised in the design of co-creative systems, but some general design principles are overlooked in the process. The principles of multidisciplinary development teams and user participation are fulfilled to some degree, but designers typically report only a few evaluation iterations, and iteration and evaluation are mostly realised at the final stages of the process. Descriptions of the early design process, such as how to investigate the creative context of use and collect ideas from users for guiding design are scarce in literature. The design of co-creative systems also seems to require some sort of adaptation of original autonomously co-creative methods to allow for the increased interaction required for successful co-creation.

Compared to creativity support systems design, human-computer co-creativity designers seem to follow a new design paradigm that accepts the system as a co-creative participant instead of focusing on the facilitation of co-creation among human creators. Design paradigms appear also to focus increasingly on autotelic and emergent co-creative activities. Design tools for human-computer co-creativity are gradually developing, with many interesting interaction paradigms, design guidelines and some design patterns suggested in recent years. This development should be encouraged by improving design tools and evaluation methodology that can support early phases of design of co-creative systems.

My second question, *how can qualitative evaluation guide the design of a co-creative system,* focuses on formative evaluation. Formative evaluation is used in both interaction design and computational creativity as a way to improve prototypes or methods. The overall quality of an evaluation method can be assessed through the three dimensions mentioned in Chapter 4: scientific quality, usefulness, and practicability. My second research question focuses on the usefulness of the evaluation results. Usefulness considers the application of evaluation data in different contexts and phases of design. In Paper II, we describe a qualitative usability evaluation of the first prototype of the Poetry Machine system. The results of our evaluation include a list of usability errors and improvement ideas from our evaluation participants. The usability errors were corrected for the next prototype version and the feedback and ideas were translated into actions for improving the system. The ideas collected from users were useful in re-designing the poetry generation algorithms, which have now improved grammaticality and semantic coherence [53] according to user feedback.

My third question, *how can quantitative evaluation be used to compare different co-creative processes in a meaningful way,* focuses on the meaningful selection of evaluation metrics for quantitative, comparative evaluation. Since the field is very young, ad-hoc metrics are often used in evaluation. This makes it difficult to assess the quality of the evaluations in practise, and to compare results across different studies. In Papers III and IV, we examine how to derive metrics for human-computer co-creativity evaluation from the related fields of computational creativity, creativity support systems, and user experience evaluation. We also examine the relationships between different metrics. In our investigation we found that in addition to the careful selection of metrics, the method of comparison also affects the usefulness of the results. By having our users complete the same creative task of writing a poem with three different methods, we were able to ask the participants also to rank their experiences with relation to each other instead of relying on separate Likert-scale estimates of the situations. This approach complements the recommendations of Yannakakis et al. [127] to evaluate emotions as relative phenomena. The quantitative ranking tool we developed can also be used to facilitate half-structured interviews with the evaluation participants to ground the metrics in their experiences.

My final question, *how can the human-computer co-creative process be described in a way that can be used to guide design decisions,* considers modelling the human-computer co-creative process. In Paper V, we describe how Wiggins' Creative Systems Framework [124] can be extended to describe human-computer co-creativity as a search. Our model describes how transformative behaviour can be used to facilitate human-computer co-creativity in meaningful ways: We identified a number of different situations in which the human and the computational collaborator struggle to produce meaningful products for each other to review. These can be translated into a number of decisions designers will need to make in planning co-creative systems. As such the model extends the tools available for designing human-computer co-creativity by approaching design decisions from a perspective familiar to computational creativity researchers. I discussed different approaches to modelling creativity in Chapter 5, and compared our model to an enactive model of creativity [30]. I found these two models complementary, and suggested using them both for designing more autonomous computational collaborators. However, additional models are needed in order to facilitate meta-level discussions between the human and the computational collaborator.

6.2 Future Work

There are multiple ways to improve the design, evaluation and modelling of human-computer co-creativity:

First, in order to improve design practise in the field, we need more design information. This requires human-computer co-creativity designers to report their design decisions and early phases of their design processes more carefully.

Second, we need more methods to evaluate conceptual work, such as interaction paradigms, or concept ideas with low fidelity prototypes. At the moment evaluation of human-computer co-creativity seems to focus on evaluation of interactive, high-fidelity prototypes, which are relatively expensive to produce and thus limit the amount of different concepts to be tested out and evaluated with real users. Here we need concrete examples of different prototyping and evaluation methods as well as comparisons of their reliability, usefulness and practicability in different contexts.

Third, we need more methodological work in considering what factors constitute a good co-creative experience. These factors are needed to construct more robust evaluation metrics for comparing and differentiating co-creative systems. Such work is also needed for conducting any type of systematic evaluation of the early prototyping methods discussed above. Here we can learn from and join forces with other interaction designers who are interested in investigating artificially intelligent interfaces.

Finally, models for human-computer co-creativity need to be extended to cover the exchange of meta-information about the creative process, products and the producers themselves. This information is essential if we are to achieve goal-oriented co-creativity in a meaningful way, or if we want to communicate affect between the collaborators. Facilitating such interactions requires considering various meta-levels of the creativity of computational collaborators. New meta-level insight would considerably improve the chances of computational collaborators being accepted as autonomous creative entities and may also be of interest to computational creativity researchers and the community of artificial intelligence researchers in general.

References

[1] Abra, J. Collaboration in creative work: An initiative for investigation. *Creativity Research Journal 7*, 1 (1994), 1–20.

[2] Amabile, T. M. A model of creativity and innovation in organizations. *Research in organizational behavior 10*, 1 (1988), 123–167.

[3] Aragon, C. R., and Williams, A. Collaborative creativity: A complex systems model with distributed affect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada, May 07 - 12, 2011), CHI '11, ACM, New York, NY, USA, pp. 1875–1884.

[4] Blackwell, A. F. The reification of metaphor as a design tool. *ACM Transactions on Computer-Human Interaction (TOCHI) 13*, 4 (2006), 490–530.

[5] Bobrow, D. G., and Brady, J. Artificial intelligence 40 years later. *Artificial Intelligence 103*, 1 (1998), 1 – 4.

[6] Boden, M. A. Creativity and artificial intelligence. *Artificial Intelligence 103*, 1 (1998), 347 – 356.

[7] Bown, O. Performer interaction and expectation with live algorithms: experiences with zamyatin. *Digital Creativity 29*, 1 (2018), 37–50.

[8] Bown, O. Empirically grounding the evaluation of creative systems: Incorporating interaction design. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, June 10-13, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., computationalcreativity.net, pp. 112–119.

[9] Bown, O. Attributing creative agency: Are we doing it right? In *Proceedings of the Sixth International Conference on Computational*

Creativity (Park City, Utah, USA, June 29 - July 2, 2015), H. Toivonen, S. Colton, M. Cook, and D. Ventura, Eds., computationalcreativity.net, pp. 17–22.

[10] Bown, O. Player responses to a live algorithm: Conceptualising computational creativity without recourse to human comparisons? In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, USA, June 29 - July 2, 2015), H. Toivonen, S. Colton, M. Cook, and D. Ventura, Eds., computationalcreativity.net, pp. 126–133.

[11] Bown, O., and Brown, A. R. Interaction design for metacreative systems. In *New Directions in Third Wave Human-Computer Interaction: Volume 1 - Technologies*, M. Filimowicz and V. Tzankova, Eds. Springer International Publishing, Cham, 2018, pp. 67–87.

[12] Bown, O., Eldridge, A., and McCormack, J. Understanding interaction in contemporary digital music: from instruments to behavioural objects. *Organised Sound 14*, 2 (2009), 188–196.

[13] Bray, L., and Bown, O. Applying core interaction design principles to computational creativity. In *Proceedings of the Seventh International Conference on Computational Creativity* (UPMC, Paris, France, June 27 - July 1, 2016), F. Pachet, A. Cardoso, V. Corruble, and F. Ghedini, Eds., Sony CSL Paris, France, pp. 93–97.

[14] Bray, L., Bown, O., and Carey, B. How can we deal with the design principle of visibility in highly encapsulated computationally creative systems ? In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, Georgia, USA, June 19-23, 2017), A. K. Goel, A. Jordanous, and A. Pease, Eds., Association for Computational Creativity (ACC), pp. 65–71.

[15] Brooke, J., et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry 189*, 194 (1996), 4–7.

[16] Brown, A. R. Creative improvisation with a reflexive musical bot. *Digital Creativity 29*, 1 (2018), 5–18.

[17] Candy, L., and Edmonds, E. Modeling co-creativity in art and technology. In *Proceedings of the 4th Conference on Creativity & Cognition* (Loughborough, UK, October 13 - 16, 2002), C&C '02, ACM, New York, NY, USA, pp. 134–141.

[18] Cardoso, A., Veale, T., and Wiggins, G. A. Converging on the divergent: The history (and future) of the international joint workshops in computational creativity. *AI Magazine 30*, 3 (2009), 15–22.

[19] Carroll, E. A. Convergence of self-report and physiological responses for evaluating creativity support tools. In *Proceedings of the 8th ACM Conference on Creativity and Cognition* (Atlanta, Georgia, USA, November 03 - 06, 2011), ACM, New York, NY, USA, pp. 455–456.

[20] Carroll, J. M. Human-computer interaction: Psychology as a science of design. *Annual Review of Psychology 48*, 1 (1997), 61–83.

[21] Cherry, E., and Latulipe, C. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction 21*, 4 (June 2014), 1–25.

[22] Clark, E., Ross, A. S., Tan, C., Ji, Y., and Smith, N. A. Creative writing with a machine in the loop: Case studies on slogans and stories. In *23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan, March 07 - 11, 2018), IUI '18, ACM, New York, NY, USA, pp. 329–340.

[23] Colton, S. The painting fool: Stories from building an automated painter. In *Computers and Creativity*, J. McCormack and M. d'Inverno, Eds. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 3–38.

[24] Colton, S. Creativity versus the perception of creativity in computational systems. In *AAAI spring symposium: creative intelligent systems* (Palo Alto, California, March 26-28, 2008), D. Ventura, M. L. Maher, and S. Colton, Eds., pp. 14–20.

[25] Colton, S., and Wiggins, G. A. Computational creativity: The final frontier? In *ECAI 2012: 20th European Conference on Artificial Intelligence* (August 27 - 31 Montellier, France, August 27 - 31, 2012), L. D. Raedt, C. Bessière, D. Dubois, P. Doherty, P. Frasconi, F. Heintz, and P. J. F. Lucas, Eds., pp. 21–26.

[26] Compton, K., and Mateas, M. Casual creators. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, USA, June 29 - July 2, 2015), H. Toivonen,

S. Colton, M. Cook, and D. Ventura, Eds., computationalcreativity.net, pp. 228–235.

[27] Corneli, J. An institutional approach to computational social creativity. In *Proceedings of the Seventh International Conference on Computational Creativity* (UPMC, Paris, France, June 27 - July 1, 2016), F. Pachet, A. Cardoso, V. Corruble, and F. Ghedini, Eds., Sony CSL Paris, France, pp. 131–138.

[28] Craveirinha, R., Barreto, N., and Roque, L. Towards a taxonomy for the clarification of pcg actors' roles. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA, October 16 - 19, 2016), CHI PLAY '16, ACM, New York, NY, USA, pp. 244–253.

[29] Davis, N. Human-computer co-creativity: Blending human and computational creativity. In *Doctoral Consortium of the Ninth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, AIIDE-13* (Boston, Massachusetts, USA, October 14-18, 2013), G. Sukthankar and I. Horswill, Eds., AAAI, pp. 9–12.

[30] Davis, N., Hsiao, C.-P., Popova, Y., and Magerko, B. An enactive model of creativity for computational collaboration and co-creation. In *Creativity in the Digital Age*, N. Zagalo and P. Branco, Eds. Springer London, London, 2015, pp. 109–133.

[31] Davis, N., Hsiao, C.-P., Yashraj Singh, K., Li, L., and Magerko, B. Empirically studying participatory sense-making in abstract drawing with a co-creative cognitive agent. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA, March 07 - 10, 2016), IUI '16, ACM, New York, NY, USA, pp. 196–207.

[32] Davis, N. M., Popova, Y., Sysoev, I., Hsiao, C., Zhang, D., and Magerko, B. Building artistic computer colleagues with an enactive model of creativity. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, June 10-13, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., computationalcreativity.net, pp. 38–45.

[33] Dearden, A., and Finlay, J. Pattern languages in hci: A critical review. *Human-Computer Interaction 21*, 1 (2006), 49–102.

[34] Deterding, S., Hook, J., Fiebrink, R., Gillies, M., Gow, J., Akten, M., Smith, G., Liapis, A., and Compton, K. Mixed-initiative creative interfaces. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA, May 06 - 11, 2017), CHI EA '17, ACM, New York, NY, USA, pp. 628–635.

[35] Dey, A. K. Understanding and using context. *Personal and Ubiquitous Computing 5*, 1 (Jan. 2001), 4–7.

[36] d'Inverno, M., and McCormack, J. Heroic versus collaborative ai for the arts. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015* (July 25–31 Buenos Aires, Argentina, July 25-31, 2015), Q. Yang and M. Wooldridge, Eds., AAAI Press, pp. 2438–2444.

[37] DiPaola, S., and Gabora, L. Incorporating characteristics of human creativity into an evolutionary art algorithm. *Genetic Programming and Evolvable Machines 10*, 2 (2009), 97–110.

[38] DiPaola, S., McCaig, G., Carlson, K., Salevati, S., and Sorenson, N. Adaptation of an autonomous creative evolutionary system for real-world design application based on creative cognition. In *Proceedings of the Fourth International Conference on Computational Creativity* (Sidney, Australia, June 12-14, 2013), M. L. Maher, T. Veale, R. Saunders, and O. Bown, Eds., pp. 40–47.

[39] Dix, A., Finlay, J., Abowd, G. D., and Beale, R. *Human-computer interaction Third edition.* Pearson Educational Limited, Essex, England, 2004.

[40] Edmonds, E. A., and Candy, L. Computer support for creativity. *International Journal of Human-Computer Studies 63*, 4 (2005), 363 – 364.

[41] Edwards, H., and Benedyk, R. A comparison of usability evaluation methods for child participants in a school setting. In *Proceedings of the 6th international conference on Interaction design and children* (Aalborg, Denmark, June 06 - 08, 2007), ACM, pp. 9–16.

[42] Eindhoven, J. E., and Vinacke, W. E. Creative processes in painting. *The Journal of General Psychology 47*, 2 (1952), 139–164.

[43] FALLMAN, D. The interaction design research triangle of design practice, design studies, and design exploration. *Design Issues 24*, 3 (2008), 4–18.

[44] FAROOQ, U., CARROLL, J. M., AND GANOE, C. H. Designing for creativity in computer-supported cooperative work. *International Journal of E-Collaboration 4*, 4 (Oct 2008), 51–75.

[45] FITZGERALD, T., GOEL, A., AND THOMAZ, A. Human-robot co-creativity: Task transfer on a spectrum of similarity. In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, Georgia, USA, June 19-23, 2017), A. K. Goel, A. Jordanous, and A. Pease, Eds., Association for Computational Creativity (ACC), pp. 104 – 111.

[46] FORTUIN, V., WEBER, R. M., SCHRIBER, S., WOTRUBA, D., AND GROSS, M. Inspireme: Learning sequence models for stories. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18)* (New Orleans, Louisiana, USA, February 2-7, 2018), S. A. McIlraith and K. Q. Weinberger, Eds., AAAI Press, pp. 2–7.

[47] GABRIEL, A., MONTICOLO, D., CAMARGO, M., AND BOURGAULT, M. Creativity support systems: A systematic mapping study. *Thinking Skills and Creativity 21* (2016), 109–122.

[48] GERVÁS, P. Dynamic inspiring sets for sustained novelty in poetry generation. In *Proceedings of the Second International Conference on Computational Creativity* (Mexico City, Mexico, April 27-29, 2011), D. Ventura, P. Gervás, D. F. Harrell, M. L. Maher, A. Pease, and G. A. Wiggins, Eds., computationalcreativity.net, pp. 111–116.

[49] GRACE, K., AND MAHER, M. L. What to expect when you're expecting: The role of unexpectedness in computationally evaluating creativity. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, June 10-13, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., computationalcreativity.net, pp. 120–128.

[50] GRACE, K., AND MAHER, M. L. Specific curiosity as a cause and consequence of transformational creativity. In *Proceedings of the Sixth*

International Conference on Computational Creativity (Park City, Utah, USA, June 29 - July 2, 2015), H. Toivonen, S. Colton, M. Cook, and D. Ventura, Eds., computationalcreativity.net, pp. 260–267.

[51] Grace, K., Maher, M. L., Davis, N., and Eltayeby, O. Surprise walks: Encouraging users towards novel concepts with sequential suggestions. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, June 25-29, 2018), F. Pachet, A. Jordanous, and C. León, Eds., Association for Computational Creativity (ACC), pp. 208–215.

[52] Gravina, D., Liapis, A., and Yannakakis, G. Surprise search: Beyond objectives and novelty. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016* (New York, NY, USA, Denver, Colorado, USA, July 20-24, 2016), GECCO '16, ACM, pp. 677–684.

[53] Hämäläinen, M. Harnessing nlg to create finnish poetry automatically. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, June 25-29, 2018), F. Pachet, A. Jordanous, and C. León, Eds., Association for Computational Creativity (ACC), pp. 9–15.

[54] Hartson, H. R., Andre, T. S., and Williges, R. C. Criteria for evaluating usability evaluation methods. *International Journal of Human-Computer Interaction 13*, 4 (2001), 373–410.

[55] Hennessey, B. A., and Amabile, T. M. Creativity. *Annual Review of Psychology 61*, 1 (2010), 569–598.

[56] Hewett, T., Czerwinski, M., Terry, M., Nunamaker, J., Candy, L., Kules, B., and Sylvan, E. Creativity support tool evaluation methods and metrics. In *National Science Foundation Workshop Report on Creativity Support Tools* (Washington, DC, June 13-14, 2005), B. Shneiderman, G. Fischer, M. Czerwinski, B. Myers, and M. Resnick, Eds., National Science Foundation, pp. 10–24.

[57] Hodhod, R., and Magerko, B. Closing the cognitive gap between humans and interactive narrative agents using shared mental models. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA, March 07 - 10, 2016), IUI '16, ACM, New York, NY, USA, pp. 135–146.

[58] Hodhod, R., and Magerko, B. Reaching cognitive consensus with improvisational agents. In *Proceedings of the Eighth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, AIIDE-12* (Stanford, California, USA, October 8-12, 2012), M. Riedl and G. Sukthankar, Eds., AAAI Press, pp. 14–19.

[59] Hourcade, J. P. Interaction design and children. *Foundations and Trends® in Human–Computer Interaction 1*, 4 (2008), 277–392.

[60] Howard, T. J., Culley, S. J., and Dekoninck, E. Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Design studies 29*, 2 (2008), 160–180.

[61] Hübscher, C., Pauwels, S. L., Roth, S. P., Bargas-Avila, J. A., and Opwis, K. The organization of interaction design pattern languages alongside the design process. *Interacting with Computers 23*, 3 (2011), 189–201.

[62] ISO/IEC. Iso 9241-210 ergonomics of human-system interaction – part 210: Human-centered design for interactive systems, 2010.

[63] Jacob, M., and Magerko, B. Interaction-based authoring for scalable co-creative agents. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, USA, June 29 - July 2, 2015), H. Toivonen, S. Colton, M. Cook, and D. Ventura, Eds., computationalcreativity.net, pp. 236–243.

[64] Jaimes, A., and Sebe, N. Multimodal human-computer interaction: A survey. *Computer Vision and Image Understanding 108*, 1 (2007), 116 – 134. Special Issue on Vision for Human-Computer Interaction.

[65] Jennings, K. E. Developing creativity: Artificial barriers in artificial intelligence. *Minds and Machines 20*, 4 (Nov 2010), 489–501.

[66] Johnson, H., and Carruthers, L. Supporting creative and reflective processes. *International Journal of Human-Computer Studies 64*, 10 (2006), 998 – 1030.

[67] Jordanous, A. A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation 4*, 3 (2012), 246–279.

[68] Jordanous, A. Four pppperspectives on computational creativity in theory and in practice. *Connection Science 28*, 2 (2016), 194–216.

[69] Jordanous, A. Co-creativity and perceptions of computational agents in co-creativity. In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, Georgia, USA, June 19-23, 2017), A. K. Goel, A. Jordanous, and A. Pease, Eds., Association for Computational Creativity (ACC), pp. 159 – 166.

[70] Jordanous, A. Stepping back to progress forwards: Setting standards for meta-evaluation of computational creativity. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, June 10-13, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., computationalcreativity.net, pp. 129–136.

[71] Kantosalo, A., and Riihiaho, S. Usability testing and feedback collection in a school context: Case poetry machine. *Ergonomics in Design 0*, 0 (2019). Preprint available online.

[72] Karimi, P., Grace, K., Maher, M. L., and Davis, N. Evaluating creativity in computational co-creative systems. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, June 25-29, 2018), F. Pachet, A. Jordanous, and C. León, Eds., Association for Computational Creativity (ACC), pp. 104–111.

[73] Khalifa, A., Barros, G. A. B., and Togelius, J. Deeptingle. In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, Georgia, USA, June 19-23, 2017), A. K. Goel, A. Jordanous, and A. Pease, Eds., Association for Computational Creativity (ACC), pp. 167 –174.

[74] Koch, J. Design implications for designing with a collaborative ai. In *AAAI Spring Symposium Series, Designing the User Experience of Machine Learning Systems* (Stanford University, March 27- 29, 2017), pp. 415–418.

[75] Lachner, F., Naegelein, P., Kowalski, R., Spann, M., and Butz, A. Quantified ux: Towards a common organizational understanding of user experience. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden, October 23 - 27, 2016), NordiCHI '16, ACM, New York, NY, USA, pp. 56:1–56:10.

[76] Lamb, C., Brown, D. G., and Clarke, C. L. A. Evaluating computational creativity: An interdisciplinary tutorial. *ACM Computing Surveys 51*, 2 (2018), 1–34.

[77] Licklider, J. C. R. Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics HFE-1*, 1 (March 1960), 4–11.

[78] Linkola, S., Kantosalo, A., Männistö, T., and Toivonen, H. Aspects of self-awareness: An anatomy of metacreative systems. In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, Georgia, USA, June 19-23, 2017), A. K. Goel, A. Jordanous, and A. Pease, Eds., Association for Computational Creativity (ACC), pp. 189 – 196.

[79] Loughran, R., and O'Neill, M. Application domains considered in computational creativity. In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, Georgia, USA, June 19-23, 2017), A. K. Goel, A. Jordanous, and A. Pease, Eds., Association for Computational Creativity (ACC), pp. 197 – 204.

[80] Lubart, T. How can computers be partners in the creative process: classification and commentary on the special issue. *International Journal of Human-Computer Studies 63*, 4 (2005), 365–369.

[81] Lubart, T. I. Models of the creative process: Past, present and future. *Creativity Research Journal 13*, 3-4 (2001), 295–308.

[82] Maher, M. L. Computational and collective creativity: Who's being creative? In *Proceedings of the Third International Conference on Computational Creativity* (Dublin, Ireland, May 30 - June 1, 2012), M. L. Maher, K. J. Hammond, A. Pease, R. P. y Pérez, D. Ventura, and G. A. Wiggins, Eds., computationalcreativity.net, pp. 67–71.

[83] Maher, M. L., Fisher, D. H., and Brady, K. Computational models of surprise as a mechanism for evaluating creative design,. In *Proceedings of the Fourth International Conference on Computational Creativity* (Sidney, Australia, June 12-14, 2013), M. L. Maher, T. Veale, R. Saunders, and O. Bown, Eds., computationalcreativity.net, pp. 147–151.

[84] Mamykina, L., Candy, L., and Edmonds, E. Collaborative creativity. *Communications of the ACM 45*, 10 (Oct. 2002), 96–99.

[85] Markopoulos, P., and Bekker, M. On the assessment of usability testing methods for children. *Interacting with computers 15*, 2 (2003), 227–243.

[86] McCormack, J., and d'Inverno, M. Designing improvisational interfaces. In *Proceedings of the Seventh International Conference on Computational Creativity* (UPMC, Paris, France, June 27 - July 1, 2016), F. Pachet, A. Cardoso, V. Corruble, and F. Ghedini, Eds., Sony CSL Paris, France, pp. 98–106.

[87] Morris, D., and Secretan, J. Computational creativity support: Using algorithms and machine learning to help people be more creative. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (Boston, MA, USA, April 04 - 09, 2009), CHI EA '09, ACM, New York, NY, USA, pp. 4733–4736.

[88] Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., and Doares, L. M. Process analytic models of creative capacities. *Creativity Research Journal 4*, 2 (1991), 91–122.

[89] Nakakoji, K. Meanings of tools, support, and uses for creative design processes. In *International Design Research Symposium 0́6, Seoul, Korea* (2006), pp. 156–165.

[90] Negrete-Yankelevich, S., and Zaragoza, N. M. The apprentice framework: planning and assessing creativity. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, June 10-13, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., computationalcreativity.net, pp. 280–283.

[91] Nemiro, J. E. The creative process in virtual teams. *Communication Research Journal 14*, 1 (2002), 69–83.

[92] Norton, D., Heath, D., and Ventura, D. Autonomously managing competing objectives to improve the creation and curation of artifacts. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, June 10-13, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., computationalcreativity.net, pp. 23–32.

[93] Norton, D., Heath, D., and Ventura, D. Autonomously creating quality images. In *Proceedings of the Second International Conference on Computational Creativity* (Mexico City, Mexico, April 27-29,

2011), D. Ventura, P. Gervás, D. F. Harrell, M. L. Maher, A. Pease, and G. A. Wiggins, Eds., computationalcreativity.net, pp. 10–15.

[94] Pérez y Pérez, R. The computational creativity continuum. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, June 25-29, 2018), F. Pachet, A. Jordanous, and C. León, Eds., Association for Computational Creativity (ACC), pp. 177–184.

[95] Read, J. C., and MacFarlane, S. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In *Proceedings of the 2006 conference on Interaction design and children* (2006), ACM, pp. 81–88.

[96] Resnick, M., Myers, B., Nakakoji, K., Shneiderman, B., Pausch, R., Selker, T., and Eisenberg, M. Design principles for tools to support creative thinking. In *National Science Foundation Workshop Report on Creativity Support Tools* (Washington, DC, June 13-14, 2005), B. Shneiderman, G. Fischer, M. Czerwinski, B. Myers, and M. Resnick, Eds., National Science Foundation, pp. 25–36.

[97] Rhodes, M. An analysis of creativity. *The Phi Delta Kappan 42*, 7 (1961), 305–310.

[98] Ritchie, G. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines 17*, 1 (2007), 67–99.

[99] Ritchie, G. Assessing creativity. In *Proceedings of the AISB symposium on artificial intelligence and creativity in arts and science* (York, England, 21st-24th March, 2001), AISB, pp. 3–11.

[100] Roemmele, M., and Gordon, A. S. Creative help: a story writing assistant. In *International Conference on Interactive Digital Storytelling. Lecture Notes in Computer Science, vol 9445* (2015), Springer, Cham, pp. 81–92.

[101] Roemmele, M., and Gordon, A. S. Automated assistance for creative writing with an rnn language model. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion* (Tokyo, Japan, March 07 - 11, 2018), IUI '18 Companion, ACM, New York, NY, USA, pp. 1–2.

[102] Rogers, Y., Sharp, H., and Preece, J. *Interaction Design: Beyond Human Computer Interaction.*, 3rd edition ed. Wiley, 2011.

[103] Saunders, R., Chee, E., and Gemeinboeck, P. Evaluating human-robot interaction with embodied creative systems. In *Proceedings of the Fourth International Conference on Computational Creativity* (Sidney, Australia, June 12-14, 2013), M. L. Maher, T. Veale, R. Saunders, and O. Bown, Eds., computationalcreativity.net, pp. 205–209.

[104] Shneiderman, B. Creating creativity: User interfaces for supporting innovation. *ACM Transactions on Computer-Human Interaction 7*, 1 (Mar. 2000), 114–138.

[105] Shneiderman, B. Creativity support tools: accelerating discovery and innovation. *Communications of the ACM 50*, 12 (2007), 20–32.

[106] Shneiderman, B., Fischer, G., Czerwinski, M., Resnick, M., and Myers, B. Introduction to workshop report. In *National Science Foundation Workshop Report on Creativity Support Tools* (Washington, DC, June 13-14, 2005), B. Shneiderman, G. Fischer, M. Czerwinski, B. Myers, and M. Resnick, Eds., National Science Foundation, pp. 4–9.

[107] Sloan, M. C. Aristotle's nicomachean ethics as the original locus for the septem circumstantiae. *Classical Philology 105*, 3 (2010), 236–251.

[108] Smith, G., Whitehead, J., and Mateas, M. Tanagra: A mixed-initiative level design tool. In *International Conference on the Foundations of Digital Games, FDG'10* (Pacific Grove, CA, USA, June 19-21, 2010), I. Horswill and Y. Pisan, Eds., ACM, pp. 209–216.

[109] Sutcliffe, A. On the effective use and reuse of hci knowledge. *ACM Transactions on Computer-Human Interaction (TOCHI) 7*, 2 (2000), 197–221.

[110] Swartjes, I., and Theune, M. Iterative authoring using story generation feedback: debugging or co-creation? In *Joint International Conference on Interactive Digital Storytelling. Lecture Notes in Computer Science, vol. 5915* (2009), I. A. Iurgel, N. Zagalo, and P. Petta, Eds., Springer, pp. 62–73.

[111] Terry, M., and Mynatt, E. D. Recognizing creative needs in user interface design. In *Proceedings of the 4th Conference on Creativity & Cognition* (Loughborough, UK, October 13-16, 2002), C&C '02, ACM, New York, NY, USA, pp. 38–44.

[112] TERVEEN, L. G. Overview of human-computer collaboration. *Knowledge-Based Systems 8*, 2-3 (June 1995), 67–81.

[113] TOIVANEN, J. M., JÄRVISALO, M., AND TOIVONEN, H. Harnessing constraint programming for poetry composition. In *Proceedings of the Fourth International Conference on Computational Creativity* (Sidney, Australia, June 12-14, 2013), M. L. Maher, T. Veale, R. Saunders, and O. Bown, Eds., computationalcreativity.net, pp. 160–167.

[114] TOIVANEN, J. M., TOIVONEN, H., VALITUTTI, A., AND GROSS, O. Corpus-based generation of content and form in poetry. In *Proceedings of the Third International Conference on Computational Creativity* (Dublin, Ireland, May 30 - June 1, 2012), M. L. Maher, K. J. Hammond, A. Pease, R. P. y Pérez, D. Ventura, and G. A. Wiggins, Eds., computationalcreativity.net, pp. 175–179.

[115] TUBB, R., AND DIXON, S. A four strategy model of creative parameter space interaction. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, June 10-13, 2014), S. Colton, D. Ventura, N. Lavrač, and M. Cook, Eds., computationalcreativity.net, pp. 16–22.

[116] VAN KESTEREN, I. E., BEKKER, M. M., VERMEEREN, A. P., AND LLOYD, P. A. Assessing usability evaluation methods on their effectiveness to elicit verbal comments from children subjects. In *Proceedings of the 2003 conference on Interaction design and children* (Preston, England, July 01 - 03, 2003), ACM, pp. 41–49.

[117] VENTURA, D. How to build a cc system. In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, Georgia, USA, June 19-23, 2017), A. K. Goel, A. Jordanous, and A. Pease, Eds., Association for Computational Creativity (ACC), pp. 253–260.

[118] VERMEEREN, A. P. O. S., LAW, E. L.-C., ROTO, V., OBRIST, M., HOONHOUT, J., AND VÄÄNÄNEN-VAINIO-MATTILA, K. User experience evaluation methods: Current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (New York, NY, USA, Reykjavik, Iceland, October 16 - 20, 2010), NordiCHI '10, ACM, pp. 521–530.

[119] WALLER, A., BLACK, R., O'MARA, D. A., PAIN, H., RITCHIE, G., AND MANURUNG, R. Evaluating the standup pun generating

software with children with cerebral palsy. *ACM Transactions on Accessible Computing 1*, 3 (2009), 1–27.

[120] Wang, K., and Nickerson, J. V. A literature review on individual creativity support systems. *Computers in Human Behavior 74* (2017), 139–151.

[121] Watanabe, K., Matsubayashi, Y., Inui, K., Nakano, T., Fukayama, S., and Goto, M. Lyrisys: An interactive support system for writing lyrics based on topic transition. In *Proceedings of the 22Nd International Conference on Intelligent User Interfaces* (New York, NY, USA, Limassol, Cyprus, March 13 - 16, 2017), IUI '17, ACM, pp. 559–563.

[122] Wen, M., Baym, N., Tamuz, O., Teevan, J., Dumais, S. T., and Kalai, A. OMG UR funny! computer-aided humor with an application to chat. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, USA, June 29 - July 2, 2015), H. Toivonen, S. Colton, M. Cook, and D. Ventura, Eds., computationalcreativity.net, pp. 86–93.

[123] Wicke, P., and Veale, T. Interview with the robot: Question-guided collaboration in a storytelling system. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, June 25-29, 2018), F. Pachet, A. Jordanous, and C. León, Eds., Association for Computational Creativity (ACC), pp. 56–63.

[124] Wiggins, G. A. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems 19*, 7 (2006), 449–458.

[125] Wiggins, G. A. Searching for computational creativity. *New Generation Computing 24*, 3 (2006), 209–222.

[126] Winston, L., and Magerko, B. Turn-taking with improvisational co-creative agents. In *Proceedings of the Thirteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-17),* (Snowbird, Little Cottonwood Canyon, Utah, USA, October 5-9, 2017), B. Magerko and J. P. Rowe, Eds., AAAI Press, pp. 129–136.

[127] YANNAKAKIS, G. N., COWIE, R., AND BUSSO, C. The ordinal nature of emotions. In *The Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)* (San Antonio, TX, USA, October 23-26, 2017), pp. 248–255.

[128] YANNAKAKIS, G. N., LIAPIS, A., AND ALEXOPOULOS, C. Mixed-initiative co-creativity. In *Proceedings of the 9th International Conference on the Foundations of Digital Games, FDG 2014* (Liberty of the Seas, Caribbean, April 3-7, 2014), M. Mateas, T. Barnes, and I. Bogost, Eds.

[129] YANNAKAKIS, G. N., AND TOGELIUS, J. *Artificial Intelligence and Games*. Springer International Publishing AG, 2018.

[130] YEE-KING, M., AND D'INVERNO, M. Experience driven design of creative systems. In *Proceedings of the Seventh International Conference on Computational Creativity* (UPMC, Paris, France, June 27 - July 1, 2016), F. Pachet, A. Cardoso, V. Corruble, and F. Ghedini, Eds., Sony CSL Paris, France, pp. 85 – 92.

[131] YOUNG, M., AND BOWN, O. Clap-along: A negotiation strategy for creative musical interaction with computational systems. In *Proceedings of the International Conference on Computational Creativity* (Lisbon, Portugal, January 7-9, 2010), D. Ventura, A. Pease, R. P. y Pérez, G. Ritchie, and T. Veale, Eds., pp. 215 – 222.

[132] ZHAO, S. Humanoid social robots as a medium of communication. *New Media & Society 8*, 3 (2006), 401–419.