Proceedings of the Eighteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment
(AIIDE 2022)

Towards AI as a Creative Colleague in Game Level Design

Tinea Larsson, Jose Font, Alberto Alvarez

Game Lab, Computer Science and Media Technology Department, Malmö University, Sweden
tinea.larsson@hotmail.com,{jose.font,alberto.alvarez}@mau.se

Abstract

In Mixed-Initiative Co-Creative tools, the human is mostly in control of what will and can be created, delegating the AI to a more suggestive role instead of a colleague in the co-creative process. Allowing more control and agency for the AI might be an interesting path in co-creative scenarios where AI could direct and take more initiative within the co-creative task. However, the relationship between AI and human designers in creative processes is delicate, as adjusting the initiative or agency of the AI can negatively affect the user experience. In this paper, different degrees of agency for the AI are explored within the Evolutionary Dungeon Designer (EDD) to further understand MI-CC tools. A user study was performed using EDD with three varying degrees of AI agency. The study highlighted elements of frustration that the human designer experiences when using the tool and the behavior in the AI that led to possible strains on the relationship. The paper concludes with the identified issues and possible solutions and suggested further research.

Introduction

Collaboration between AI and humans to co-design and co-create content is a significant challenge and the main focus of Mixed-Initiative Co-Creativity (MI-CC), which is the joint effort by a human user and AI to create content together (Yannakakis, Liapis, and Alexopoulos 2014; Liapis et al. 2016). In an MI-CC environment, designers can unleash their creativity while the computer ensures playability, measures quality, and potentially inspires them towards more creative designs. These systems' objectives are to foster creativity and provide seamless proactive collaboration, ultimately enabling a mutually beneficial collaboration. The AI role has been categorized depending on the computer agency and initiative: nanny, pen-pal, coach, and colleague (Lubart 2005). For an AI to be a colleague, it would have to intervene in the human process and take initiatives directly affecting the end product and creative process.

Morai Maker is an AI-driven Level Editor for Super Mario Bros-style games (Guzdial et al. 2019), which aims at having an AI as a colleague, with an equal role as the human designer, both adapting to each other. The Evolutionary Dungeon Designer (EDD) is a mixed-initiative design tool to create adventure and dungeon crawler games. EDD uses an evolutionary algorithm (MAP-Elites) to constantly generate finished rooms for the user to pick and replace their design based on the user's manual designs. The AI does not have any definitive control over the design decisions. Rather it suggests content adapted to the designer's current design, and the designer has the option not to incorporate the AI in their creations (Alvarez et al. 2019). Nevertheless, it seems relevant to explore how other degrees of AI agency could affect the resulting co-creative process in terms of frustration, constraints, efficiency, or diversity, compared to when two humans create together. This comes with potential issues derived from altering the AI's agency; that human creativity can be dampened by restrictions in the creative process (Yannakakis, Liapis, and Alexopoulos 2014).

This paper explores how AI with varying degrees of agency affects the human users' design process in EDD. Three different versions of the tool are developed with varying degrees of the AI's control over the design process. These versions are then examined in a user study, and the results are analyzed to understand further the colleague relationship between humans and AI in MI-CC systems. The study also analyzes the degree of support these three AI companions have on lateral thinking, which is a vital part of the creative process. By assessing the three variants of agency, it is possible to compare the differences in the resulting creative relationships between the designer and AI, identifying factors that affect the designer's creative process in terms of frustrating elements, perceived limitations, and adaptation to their creative colleague.

Related Work

MI-CC focuses on tackling tasks between humans and AI with proactive initiative, where AI does not only assist humans but could also collaborate with them, leveraging on both their strengths (Yannakakis, Liapis, and Alexopoulos 2014; Allen 1999). *Initiative* is a multi-factor model combining: choosing the task, the agent in control and how the interaction is established, and the expected outcome (Novick and Sutton 1997). In this work, both humans and AI have the same task, and the interaction is established as turn-based, each taking discrete control. The outcome is expected to vary as AI agency increases since larger constraints are added for the human that might need to adapt towards those.

Version properties	AIv1	AIv2	AIv3
Can override AI-tiles	✓	✓	
Can override human-tiles	✓	✓	✓
Places directly		✓	✓

Table 1: Properties and differences between AI versions

Some MI-CC systems enable different collaborative approaches, which are considered in this paper. Tanagra (Smith, Whitehead, and Mateas 2011) is a design tool for platform levels where the system takes as input and constraints the current user's design and creates content fulfilling gaps around it. Morai Maker (Guzdial, Liao, and Riedl 2018) is an MI-CC tool where the human designer and AI take turns to design Super Mario Bros. levels. The AI adds content in its turn, which can be maintained or erased by the human designer, which the AI learns to adapt to through reinforcement learning. Furthermore, Lode Encoder (Bhaumik, Khalifa, and Togelius 2021) explores a creative collaboration where the human is constrained by only being able to use AI-generated content, which they need to choose to compose their design. This shows an unusual collaboration that users expressed as a playful, game-like creative process.

This paper uses EDD as the tool to explore AI agency and control. EDD is an MI-CC system where designers can create interconnected rooms composing a dungeon (Alvarez et al. 2019). As designers create their content, the AI constantly suggests content adapted to the designer's design using the Interactive Constrained MAP-Elites (IC MAP-Elites). We make extensive use of IC MAP-Elites to generate rooms that are adapted to the target room. In (Alvarez et al. 2020), the authors show that IC MAP-Elites can generate high-performing and diverse rooms from different targets and using different dimension combinations. Its adaptiveness and stability, two necessary properties, were assessed with continuously edited rooms in (Alvarez et al. 2021), showing that the designer has a positive effect and can steer the algorithm with their design.

AI Roles and Adaptability

Lubart discusses four different roles a computer might take to promote creativity; *computer as nanny*: management of creative work; *computer as pen-pal*: communication service between collaborators; *computer as coach*: Using creative enhancement techniques; and *computer as colleague*: partnership between computer and humans (Lubart 2005). This is further explored by Guzdial et al. where designers perceived the AI collaborator with more or less value depending on their desired role for the AI, varying between: *friend*, *collaborator*, *student*, or *manager* (Guzdial et al. 2019).

Establishing different roles such as colleague and collaborator might require some user model within the system. Designer modeling, as defined by Liapis et al. (Liapis, Yannakakis, and Togelius 2013), is a way to classify and predict a designer's style, goals, preferences, and processes. Preference models (Alvarez and Font 2020; Liapis, Yannakakis, and Togelius 2012) have been built based on designers' choices and used as surrogate models to evaluate further generated content. Similarly, using the designers' creation, the designers' processes and styles could be modeled to inform other systems and adapt the generated content (Liapis, Yannakakis, and Togelius 2014; Alvarez, Font, and Togelius 2022; Halina and Guzdial 2022).

Altering Human-AI Collaboration Dynamics

The standard version of EDD presents a unidirectional relationship between the human designer and the AI, where the AI can only suggest content adapted to the designer's room, called the target room. As concluded in Morai Maker (Guzdial, Liao, and Riedl 2018), allowing the human designer and the computer-controlled agent to take turns in the creative process enables the possibility of an even influence between the co-creators; thus, the human user and the AI take turns placing down tiles. Here we present three modified versions of EDD that implement three different dynamics for the human-AI co-creative process.

AI Version 1 (AIv1) - Low Degree of Agency

In this version the AI takes a suggestion colleague role. As the human manually designs a target room, the AI suggests tiles directly on top of the design. The human has the option to make use of the suggested tiles at will, placing them on the target room by clicking on them in the user interface. Both the human and the AI can override each other tiles.

AI Version 2 (AIv2) - Medium Degree of Agency

In this version, the AI places directly its recommended tiles rather than suggesting. Like in AIv1, both the human and the AI can override each other tiles.

AI Version 3 (AIv3) - High Degree of Agency

Unlike in the other versions, in AIv3, the human designer cannot be the sole contributor to the room designs. The AI places the tiles on their turn rather than suggesting, and the human cannot overwrite them. However, the AI can overwrite human tiles in their turn. This allows the exploration of how human designers react to being in a co-creative relationship where the AI has more control than them and add constraints to their design and goals.

The Design of the AI

The goal is for the AI to be perceived as dynamic, responsive, and helpful but not totally predictable. These qualities were selected to try to create a co-creator that supports the design choices of the human but also introduces unexpected elements to stimulate lateral thinking. By making the AI dynamic and responsive, the aim is to minimize the risk of unsatisfactory asymmetric design between the AI and the human, as reported by (Guzdial, Liao, and Riedl 2018), where some of the critiques mentioned that the AI was designing its own parts of the level instead of creating consistency with the human designer's contributions.

For all three versions of the AI, the AI component creates and regularly updates a list of generated tiles. IC MAP-Elites constantly runs in the background, maintaining a list of elites across its seven dimensions. When the human ends their

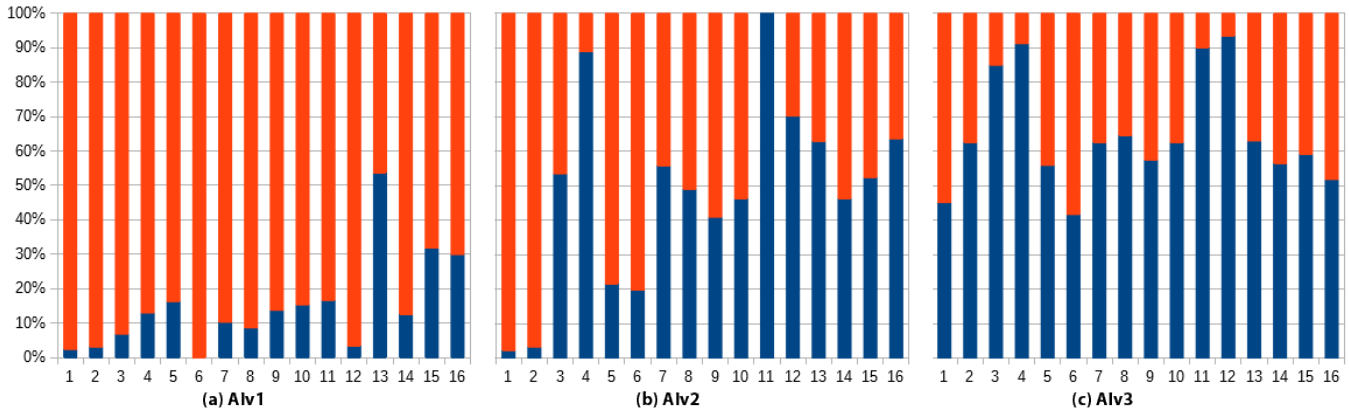

Figure 1: Percentage of tiles per co-creator in the rooms. Blue and Red bars relate to AI and Human placed tiles, respectively.

turn, the elites are processed by a KNN algorithm (K=20) that picks the set of tiles that will be used by the AI contributor comparing the elites to the target room on the seven dimensions. The resulting list of elites is further processed tile by tile, creating a final list that contains the most reoccurring tile types per position in the contribution area, which constitutes the list of generated tiles for the computer to use. From this list, and to further the perception of the AI being dynamic, it selects a random amount of tiles between 50% and 100% of the amount of human-placed tiles. However, through this process it is likely that the agent will not be human-like. For example, humans are likely to favor symmetry in a room, but the AI in this tool will consider all dimensions equal. The AI will also consider all types of tiles equally when calculating the most common tile in a position in the generated rooms.

Furthermore, depending on the level of agency, these tiles are just displayed as suggestions on the UI or directly placed when the computer takes control of the creative process. Both creators have a maximum of 12 tiles that they can contribute per turn. This was determined through experimentation during development since it is large enough to allow the designer to create a small subsection during their turn. If the human does not contribute one round before pressing ”End Turn,” the previous turn’s amount of tiles are used for the calculation to enable the human to press ”End Turn” repeatedly to let the AI keep contributing if that is desired. The available locations for the AI to contribute each turn are limited to a rectangular area surrounding the tiles the human designer recently placed, including a margin of 1 tile.

Experiment Setup

We conducted a user study to explore the user experience of using different levels of AI agency, the different design characteristics, and the relationship between the human designer and the AI. We collected both quantitative data on the AI’s impact on the co-designed end product and qualitative data through think-a-loud and semi-structured interviews regarding the users’ experience when interacting with the AI. The interview structure is inspired by the pyramid model, meaning the interviews will begin with specific questions, and gradually have more open questions, which naturally allows for a discussion towards the end. This model is chosen to support the variation of subjects the interview is desired to cover, as well as support natural transitions between the questions and their openness. The questions and user study procedure can be found in Appendix A.

Eight participants tested our tool with game design and level design experience. One participant was a professional game designer with eight years of professional experience (first participant), and seven participants were third-year Game Development students. They all had an individual digital session, where we shared our screen, and they took remote control to conduct the study. Participants accepted to participate, signed consent forms, and then received a short introduction describing the experiment and its steps. The participants were then asked to design two contiguous rooms in a dungeon, repeating this process for each of the AI variants and expressing their design decisions verbally whenever they felt like it. After using the tool, the participants were interviewed, focusing on and covering an overarching understanding of the user experience, particularly in terms of creativity and interaction with the AI.

For all the sessions, human designers could place up to 12 tiles, and the AI could place as many tiles as the human placed. The AI could contribute only in a rectangular area surrounding the tiles the human designer recently contributed with, including a margin of 1 tile. This choice is made to support a responsive and collaborative behavior of the AI that builds on the human designer’s contribution.

Results

We present the results from the user study regarding the designed rooms, the room design process with the AI, and the participant’s interview responses. Figure 1 and 2 show the tile contribution for both human and AI per room and tested version, and the final tile distribution across all versions, respectively. Figure 3 shows a sample of the designed rooms corresponding to the different AI versions, respectively.

The designed rooms, when interacting with the AIv1 (fig. 3.a), generally include a vast majority of human-placed tiles or unedited tiles (86%). Many of the rooms display trea-

	AIv1	AIv2	AIv3
Leniency	0.56±0.07	0.62±0.09	0.57±0.08
Linearity	0.91±0.02	0.92±0.02	0.91±0.01
MesoPat	0.15±0.05	0.13±0.07	0.12±0.05
SpatialPat	0.35±0.1	0.41±0.11	0.34±0.09
Symmetry	0.43±0.11	0.35±0.18	0.35±0.12
W_{dens}	0.27±0.09	0.26±0.08	0.21±0.05
W_{spar}	0.21±0.05	0.19±0.03	0.15±0.01
E_{dens}	0.24±0.07	0.27±0.06	0.3±0.06
E_{spar}	0.22±0.05	0.32±0.05	0.35±0.06
T_{dens}	0.37±0.13	0.28±0.09	0.34±0.07
T_{spar}	0.36±0.11	0.3±0.1	0.37±0.05
Steps	39.25±6.38	84.31±14.85	76.75±17.02

Table 2: Summary of the created rooms filtered by the AI version used. All values are the average of all the created rooms using the specific AI version. The first five values relates to the MAP-Elites dimensions, then the fitness of the rooms, the density and sparsity values for wall (W), enemies (E), and treasures (T), and finally the avg. steps taken to design a room.

sures and enemies placed close to each other, often with an enemy blocking a treasure. As displayed in fig. 3.b (AIv2), rooms contain more AI edited tiles (51%), and long continuous walls are less common compared to when using AIv1. Fig. 3.c shows a sample of the designs using AIv3. The resulting rooms are less symmetric, contain fewer continuous walls, and are, in general, somewhat less organized.

The relation between the total amount of tiles placed by the human designer and the AI for each version provides insight into how much the human designers incorporated the AI in their design process. When using AIv1, the human designers generally contributed with a majority of tiles to the resulting room (86% in general). When using AIv2, the results vary more between rooms and/or designers, such as fig. 3.b 1 and 9 with a general 49% of tiles placed by humans. In AIv3, the AI generally had a majority of tiles in the resulting rooms (32% placed tiles by human designers in general). Figure 2 shows the tiles by each co-creator. Whereas human designers focused mostly on walls and structures, the AI focused on removing tiles and adding floors.

Furthermore, table 2 and 3 summarizes the created rooms by AI version. Table 2 shows that there is not much variation regarding the IC MAP-Elites dimensions, but Symmetry and Meso patterns reduce as the AI gains more agency, and there are fewer spatial patterns in the high and low agency. It also shows that wall and enemy density and sparsity go lower and higher, respectively. This is expected, given that as the AI is more dominant in the design (fig. 1), walls are diminished (fig. 2), which can also be observed in the examples in fig. 3. Finally, designers spent less time on their design for AIv1 and AIv3. In AIv1, designers can edit rooms towards their goal and the AI does not modify their design (unless wanted), which means that the designer does not need to modify the room much. This is expected to increase for AIv2 and AIv3 because the AI now adds its editions, which count towards the steps. However, since the AI can add as many

	AI replaced	Human replaced	Interactions
AIv1	3.31±2.1	1.63±1.47	3.88±0.59
AIv2	24.94±6.67	12.06±5.29	8.56±3.28
AIv3	19.25±5.13	0	9.13±3.64

Table 3: Summary of the interactions between AI and human. The first two columns relate to the avg. amount of tiles that the AI or human replaced of each other. Interactions refer to the avg. number of times the human pressed "End Turn."

tiles as the designer, the designer is still doing roughly half of the steps. When using AIv3, rooms take slightly fewer steps than AIv2, probably due to the AI taking over areas where to design and designers trying to work around constraints further discussed in the next section.

Table 3 shows the interaction between human and AI. As expected, the human designers interacted less with the AI in AIv1, and in general, they overwrote less since they could decide which elements from the AI to include. Five out of the 16 rooms included AI-placed tiles where the human designer did not replace any of the AI tiles. Interactions and human and AI tile replacements increase for AIv2 and AIv3, which is also expected. The AI overwrote an avg. of 24.94 and 19.25 tiles in each subsequent version, and human designers overwrote an avg. of 12.06 in AIv2. This shows that the AI had more involvement and changed the design more often. The human designer could either use their turn to replace the tiles or continue using those tiles. However, in AIv2, the designers did not replace all of the AI tiles; instead, they also chose to work around them and replaced them when needed, allowing the AI to participate more. In 13 out of the 16 rooms created the designer replaced fewer tiles than the AI did, and for 9 of those it was less than 50%.

Perceptions of the Different AI Versions

Four participants preferred to use AIv1 due to higher controllability over the final design. Two participants preferred AIv2 since they liked the efficiency of the AI placing down tiles, but they still remained in control over the design process. One participant said that they preferred both AIv1 and AIv3, as they felt that they fulfilled different purposes. AIv1

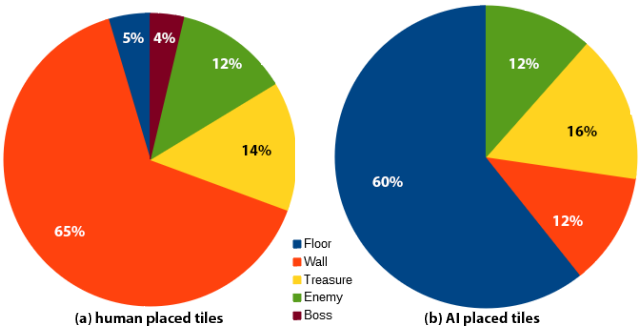

Figure 2: The types of tiles placed and their percentage of occurrence by human designers (a) and the AI (b)

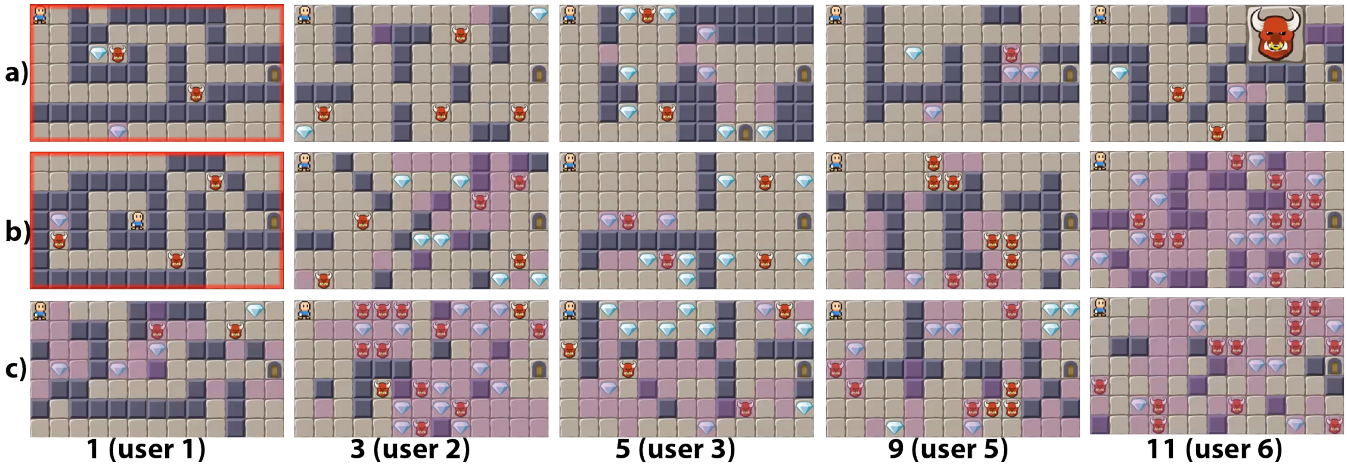

Figure 3: Sample resulting rooms from the study. a, b, and c are samples from the same users using the different versions (AIv1, AIv2, and AIv3, respectively). Red borders indicate that the room contains unreachable tiles. Purple tinted tiles are the ones added by the AI.

contributed with inspiration to get out of a writers block kind-of situation in level design, while AIv3 offered a more unusual creative experience, and created unique levels. One participant preferred AIv3 categorizing it as more efficient. However, AIv3 was the most disliked due to most participants feeling constrained by the AI's decisions, which forced themto workaround the AI's design; too invasive on their design process, and in general, feeling the AI dismissed their contributions. Two participants disliked AIv2 the most due to frustration and slowing their design process because of the repetitive work of recreating their original idea as the AI would overwrite their tiles. One participant disliked AIv1 as it was the slowest to work with.

AI's Behavior

Five participants described the AI's behavior as random and unpredictable. Many mentioned that the AI often placed floor tiles over the human-placed walls, treasures, enemies, and bosses. Additionally, two participants mentioned that the AI often broke down sub-rooms or long walls with floor tiles. The more positive descriptions of the AI's behavior were that the AI repaired their level if it had unreachable tiles by placing floor tiles on the positions wall tiles were, making all tiles reachable. One participant expressed that they thought the AI created unique-looking high-quality levels.

The Creative Relationship

Five participants expressed that the AI contributed with ideas they liked and either kept or considered incorporating. Many also expressed that they found it frustrating that the AI overwrote the human placed walls, treasures, enemies, or boss tiles with floor tiles, as they perceived this as the AI removing their contributions without adding anything new. All participants answered that they did adapt to AIv2 and AIv3 during their design process. When using AIv2, the participants adapted to the AI by either being inspired by the AI's contribution or by getting frustrated with the AI repeatedly placing tiles the human designer did not like. When using AIv3, they felt they had to adjust to whatever the AI contributed and felt forced to adapt. Additionally, six participants answered that they did not feel that the AI adapted to them. One answered that it did feel that the AI adapted to them, and one participant had no opinion.

Four participants described the relationship as working together with someone who only says no to your ideas without contributing to new ideas. Three participants described it as frustrating and/or a fight for control between the AI and the human designer. Two participants described it as an iterative collaboration and compared it to two people working on the same product but in different steps. One participant described the relationship as a brainstorming session where you work in a "Yes, and..." fashion, meaning you work iteratively and only adds to the idea and never decline the other collaborator's contribution.

The Creative Process

Four participants mentioned that the AI negatively affected their creative process, especially when the AI placed floor tiles over their placed walls, treasures, enemies, and bosses. This was described by multiple participants as the AI "destroying their creations", and many felt forced to let the AI "take control over the room." Two participants described the creative process as the AI and the human designer working against each other. Two other participants described it as a process where the AI brought forward ideas that they found interesting. One participant described it as letting the AI form an idea that the human designer finally polished.

Constraints and Design Goals

Five participants answered they felt progressively more constrained as the AI gained more control, although one participant answered that they did not feel constrained at all with any of the versions. Six participants answered they had a design goal of creating a boss room. However, the AI placed

floor tiles over the boss tiles, making this impossible in AIv2 and AIv3, forcing them to change their design goals. Two participants answered that they did not have set design goals and solely created content each turn with no specific concept for the whole room in mind.

Discussion

Willingness to Include the AI in the Design Process

All participants expressed an interest and willingness to see what the AI could come up with to design the rooms, emphasizing that they either considered or incorporated the ideas brought forward by the AI, which further supports other MI-CC research conclusions (Guzdial et al. 2019; Bhaumik, Khalifa, and Togelius 2021). However, many participants expressed multiple frustrating factors, and based on figure 1.a, when given the opportunity, the human designers did not include much of the AI's contribution but might have provided some ideas that either influenced or were part of the final design. Figure 1 also displays that as the AI got more agency, the AI tiles at the end design increased. This can be due to frustration expressed by participants where they didn't agree with the AI design and ended up handing over the control to the AI and giving up, to some extent, their aspiration to create. The lower meso- and spatial-patterns combined with the lower symmetry in AIv3 are also part of the issue. The fewer patterns that exist mean that the rooms are less structured, which, combined with no way to correct these, could result in designers feeling that the levels are more "random".

Variants of the Users

Most participants used the tool similarly except for participants 1 and 6. Participant 1 didn't want to incorporate the AI's contributions, as it can be seen in fig. 1, explicitly stating that "... I don't think level design is a good place for an AI that has more control than the human... The little details that I love in level design would never be created by an AI. Nice little references, or easter-eggs, or how humans get inspired by simple things..." On the other hand, participant 6 recurrently incorporated many of the AI's tiles. When using AIv2 and AIv3, they pressed "End Turn" repeatedly to find out what the AI would be able to create, commenting "I want to see if it can create something cool." However, while their approach was completely different, they both agreed that they would prefer the AI to create a complete room, and they could polish it from there.

Frustrating Factors and Constraints

The participants expressed multiple frustrating factors within the tool. The main point was the repetitive behavior of overwriting the human tiles with floor tiles, removing their ideas without contributing with anything of value, and the human designer feeling forced to move on from those positions and contribute somewhere else in the room. This was exacerbated when using AIv2 as, unlike AIv1, the AI placed down the tiles rather than suggesting, and unlike AIv3, the human designer still had the option to overwrite the AI-placed tiles. The human designers assign value to each tile type as they provide different aspects in level design; for instance, participants often placed enemies and treasures close to each other, possibly to create a risk and reward in the level. Figure 4 shows one example of a participant creating a room using AIv2, step by step. The human designer's first contribution includes a long continuous wall and a boss. When the AI has its first turn, it contributes with floor tiles overwriting the designer's tiles. Towards the end of the design, the human designer places a boss tile in the bottom left corner that the AI overwrites with floor tiles twice before the human designer gives up and finishes the room without a boss. This sample creation process shows that the AI tried to steer the room towards more leniency and open areas, which contradicted the human's goal.

Another main frustrating factor was the loss of control experienced by human designers when co-creating with the AI, especially AIv3. Participants expressed that the AI's decisions limited them and were forced to work around what the AI designed. As the AI gained control, they felt their creative process got increasingly constrained. This aligns with the Lode Encoder study (Bhaumik, Khalifa, and Togelius 2021), where the participants expressed frustrations with completing a playable level, as they were forced to rely on the AI to generate the option they wanted in the final stages of the level creation. Further, when using AIv3, the number of positions that the human has available decreases with every turn, while the AI can continue to place tiles on any position, which unavoidably limits the human's control over the final design.

Most of the participants felt frustrated and constrained as the AI gained more control over the design process. Additionally, all participants suggested removing the turns and constraints of the number of tiles per turn to improve the tool. Three of the eight participants expressed that adjusting the AI's role to one of an assistant to the human designer would improve their creative experience.

The Concept of a Well Performing, High Agency Co-Creator

Most participants showed a willingness to incorporate AI into the creative process, contributing with new ideas or performing services such as ensuring feasibility. Yet they were reluctant to incorporate higher agency AI, which suggests that the AI needs to be aligned with their goals, intentions, and procedures, i.e., have an accurate designer model (Liapis, Yannakakis, and Togelius 2013). Within EDD and level design tools, multiple practical improvements are to be made. Five participants described the AI's behavior as random and unpredictable, especially when overwriting human-made structures and contributions. The AI currently calculates the most common tiles in the positions of the contribution area and contributes with the tiles of highest occurrence among a set of generated elites. This contradicts how human designers perceive the design importance of tiles, valuing higher usable tiles rather than floors. The AI could then weigh higher those and the combined structures they create. Additionally, the AI could favor unedited areas before overriding human-placed tiles to support rather than override.

Another important point is that all AI versions are static

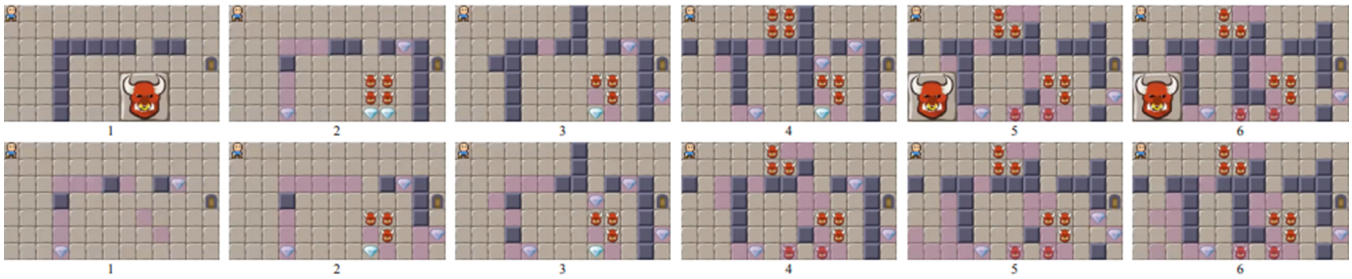

Figure 4: Sample step by step process when using AIv2. Human's turn is top row, while the AI's turn is the bottom row.

in the design process, which means that the AI follows the same procedure regardless of the agency level. The collaboration do change in the design process (e.g., suggest or directly placing tiles, or if AI tiles can be removed) but other aspects and parameters do not change or adapt to designers. These parameters are connected to the overarching design of the AI rather than the AI's agency, which might have affected the designers' perception. For instance, given that the AIv3 tiles could not be replaced, changing the amount of tiles, rectangular area, or its adaptability in regards to what the designer had created thus far could be beneficial.

Furthermore, the AI seemed to break apart walls and open up sub-rooms. This is possibly a result of the Linearity and Meso-Pattern dimensions in the MAP-Elites algorithm. The resulting elites of the generated rooms with the highest linearity will have the highest amount of traceable paths. Many participants seemed to want to create rooms with long walls, sub-rooms, paths that required the player to encounter enemies, and common aesthetical features such as symmetry. Analyzing the path designers are taking in these dimensions could better inform the search for content and the generation of elites, so the content is adapted to those preferences. However, adapting these dimensions might be counterproductive for the other dimensions, as symmetric rooms might not create balanced rooms regarding the Leniency dimension as it might not be considered as important. Another approach could be to incorporate designer modeling. By identifying possible design goals or design styles of the human co-creator, the AI can adjust its decisions and behavior to offer different levels of support depending on the human designer's behavior (Liapis, Yannakakis, and Togelius 2013; Alvarez, Font, and Togelius 2022).

Conclusion

This study explored the limitations and possibilities of an MI-CC-tool with an AI with a varied agency. We aimed at doing an initial exploratory study with static parameters, resulting in baselines to analyze what can be done and how designers experienced the system. This, in turn, opens up and continues the discussion towards AI collaborating as a colleague and enabling alternative ways to foster creativity (e.g., constraining the design space such as in (Bhaumik, Khalifa, and Togelius 2021)). Our study showed that AI gaining control over the design results in frustration and feeling constrained. Constraints are not bad per se, as they can be a way to foster creativity (Boden 2004; Acar, Tarakci, and Knippenberg 2019), but they need to be placed in a way that the human designer might feel inspired, motivated, or supported to continue the design. Human designers had to adapt towards those imposed goals instead of the other way around, which creates an unwanted dynamic when human designers perceive the AI's behavior as erratic, random, and without clear objectives.

Many of the results pointed to a general preference for an AI with a more supportive role in collaborative tools. One approach could be to have a hybrid model between what is presented in this paper and other typical MI-CC systems that focus more on suggesting final designs. The AI could take parameters from the human designer, such as an area in a room, amount of tiles, or an attribute that the human designer would like to increase in the room, but still maintain their design, effectively constraining the AI to find creative ways to achieve its goals. In EDD, designers can lock tiles to not be changed by the AI, which is something to be experimented with. Although this would give the human designer a slightly higher degree of influence on the end product compared to the AI, the constraints of how many tiles can be locked, or possibly what types of tiles can be locked, can be experimented with to adjust the relationship between AI and human designer. Currently, the search is steered, to some extent, by the designers' design, but in future work, we could bias the search even more towards interesting areas based on the creation process and the trajectories they are taking in behavior dimension space.

Additionally, using designer models is a feasible approach. By predicting design goals, adapting to phases of the design process, or identifying certain design styles and adapting to the human designer, a responsive and adaptive, intelligent, and human-like artificial co-creator could be developed. This could allow for an AI that adapts to the human designer and performs well enough that the frustrations and feelings of constraints are minimal or perceived as less prevalent as the designs turn out more similar to what the human desired.

Appendix A

Interview Procedure and Questions

Setting up a user study session

1. The conductor starts a meeting in Zoom.
2. The conductor explains the steps to the participant.

3. When consent of recording the session is acquired by the conductor, the conductor starts recording.
4. The conductor starts the tool and shares the screen.
5. The conductor allows the participant to control the conductors machine via Zoom remote control.
6. The conductor instructs the participant to perform the test.

Instructions The task is to design at least two rooms in a dungeon world, with each variant of the AI.

- Step 1: Choose the LOW level of the AI and click "Create World".
- Step 2: You are now in the World Editing View. Edit as you please. To enter the room editing view, double click the room you wish to edit.
- Step 3: You are now in the Room Editing View. Use the brushes on the left to edit the room. Click "End Turn" to end your turn, and let the AI contribute.
  - AI-placed tiles are tinted purple. AI-suggested tiles are tinted green.
  - If you are in the LOW variation of AI, click the green suggestions you'd like to place. Click continue when you want to have your turn again.
  - To go back to the World Editing View, click "Go To World View".
- Step 4: When you feel satisfied with your creation, tell me so. Restart the program. Start over at step 1 for the next AI version until all three are used once.

Interview Questions

- Q1: Which of the three versions of AI did you prefer? Why?
- Q2: Which of the three versions of AI did you find least appealing? Why?
- Q3: How would you describe the creative experience?
- Q4: What is your perception of the AI's behaviour?
- Q5: Did you feel your creativity was constrained when using any of the three AIs?
- Q6: Did you adapt to the different AI versions? In what ways?
- Q7: Did you perceive that the AI adapted to you? In what ways?
- Q8: How would you describe the relationship between designer and the AI?
- Q9: How did the AI's decisions affect your creative process?
- Q10: How did the different versions affect your design goals?
- Q11: What do you think is missing or needs to be improved for an AI as the one of the HIGH-version (with high initiative) to be used in collaborative tools?

References

Acar, O. A.; Tarakci, M.; and Knippenberg, D. v. 2019. Creativity and Innovation Under Constraints: A Cross-Disciplinary Integrative Review. *Journal of Management*, 45(1): 96–121. _eprint: https://doi.org/10.1177/0149206318805832.

Allen, J. F. 1999. Mixed-Initiative Interaction. In *Intelligent Systems, Trends and Controversies section*, 14–16.

Alvarez, A.; Dahlskog, S.; Font, J.; and Togelius, J. 2019. Empowering Quality Diversity in Dungeon Design with Interactive Constrained MAP-Elites. In *2019 IEEE Conference on Games (CoG)*.

Alvarez, A.; and Font, J. 2020. Learning the Designer's Preferences to Drive Evolution. In Castillo, P. A.; Jiménez Laredo, J. L.; and Fernández de Vega, F., eds., *Applications of Evolutionary Computation*, volume 12104 of *Lecture Notes in Computer Science*, 431–445. Cham: Springer International Publishing. ISBN 978-3-030-43722-0.

Alvarez, A.; Font, J.; Dahlskog, S.; and Togelius, J. 2020. Interactive Constrained MAP-Elites: Analysis and Evaluation of the Expressiveness of the Feature Dimensions. *IEEE Transactions on Games*, 1–1.

Alvarez, A.; Font, J.; Dahlskog, S.; and Togelius, J. 2021. Assessing the Effects of Interacting with MAP-Elites. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 17, 124–131. AAAI. Section: Poster Papers.

Alvarez, A.; Font, J.; and Togelius, J. 2022. Designer Modeling through Design Style Clustering. *IEEE Transactions on Games*, 1–1.

Bhaumik, D.; Khalifa, A.; and Togelius, J. 2021. Lode Encoder: AI-constrained co-creativity. In *2021 IEEE Conference on Games (CoG)*, 01–08. IEEE. ISBN 978-1-66543-886-5.

Boden, M. 2004. *The Creative Mind: Myths and Mechanisms*. Routledge. ISBN 978-0-415-31452-7.

Guzdial, M.; Liao, N.; Chen, J.; Chen, S.-Y.; Shah, S.; Shah, V.; Reno, J.; Smith, G.; and Riedl, M. O. 2019. Friend, Collaborator, Student, Manager: How Design of an AI-Driven Game Level Editor Affects Creators. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-5970-2. Event-place: Glasgow, Scotland Uk.

Guzdial, M.; Liao, N.; and Riedl, M. 2018. Co-Creative Level Design via Machine Learning. In *Joint Proceedings of the AIIDE 2018 Workshops co-located with 14th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2018), Edmonton, Canada, November 13-14, 2018.*

Halina, E.; and Guzdial, M. 2022. Threshold Designer Adaptation: Improved Adaptation for Designers in Co-creative Systems. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22), Special Track on AI, the Arts and Creativity*.

Liapis, A.; Yannakakis, G. N.; Alexopoulos, C.; and Lopes, P. 2016. Can Computers Foster Human Users' Creativity? Theory and Praxis of Mixed-Initiative Co-Creativity. *Digital Culture & Education*, 8(2): 136–153.

Liapis, A.; Yannakakis, G. N.; and Togelius, J. 2012. Adapting Models of Visual Aesthetics for Personalized Content Creation. *IEEE Transactions on Computational Intelligence and AI in Games*, 4(3): 213–228.

Liapis, A.; Yannakakis, G. N.; and Togelius, J. 2013. Designer modeling for personalized game content creation tools. In *Artificial Intelligence and Game Aesthetics - Papers from the 2013 AIIDE Workshop, Technical Report*, volume WS-13-19, 11–16. AI Access Foundation. ISBN 978-1-57735-634-9.

Liapis, A.; Yannakakis, G. N.; and Togelius, J. 2014. Designer modeling for Sentient Sketchbook. In *Proc. IEEE Conf. Computational Intelligence and Games*, 1–8. ISSN: 2325-4289.

Lubart, T. 2005. How can computers be partners in the creative process: Classification and commentary on the Special Issue. *International Journal of Human-Computer Studies*, 63(4): 365 – 369.

Novick, D.; and Sutton, S. 1997. What is Mixed-Initiative Interaction? In *Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction*.

Smith, G.; Whitehead, J.; and Mateas, M. 2011. Tanagra: Reactive Planning and Constraint Solving for Mixed-Initiative Level Design. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(3).

Yannakakis, G. N.; Liapis, A.; and Alexopoulos, C. 2014. Mixed-Initiative Co-Creativity. In *Proceedings of the 9th Conference on the Foundations of Digital Games*.