A Social Robot to Support Children's Role-play

MIGUEL AZINHEIRA, Institute Superior Técnico, Portugal

Creativity is a non-static trait that can be trained and improved from an early age. Children when they join the education system, have a decline in creativity due to the pressures of fitting in with their school peers. This decline can be countered with the help of creativity training tools, but such tools still look like tests and are not appealing to children. Robots have proven to be valuable tools to increase the engagement of children in activities, and when combined with storytelling have been demonstrated to help train creativity. Robots also have benefits when used in a shared autonomy context increasing users' performance in tasks while still retaining their trust. Shared autonomy allows the combination of human attributes with the attributes of a robot. In this paper, we present a study using a semi-autonomous robot that explores a simulated shared autonomy by a wizard in a storytelling activity. We measure in a storytelling activity, the impact of two different reactive strategies on creativity. These reactive strategies consisted of six autonomous behaviors that were displayed either randomly at key moments of the story or displayed according to the context and content of the story. We show that when children have lower baseline values of creativity, the semi-autonomous robot, combined with random emotional behaviors, has a positive impact on the fluency and originality of the stories.

Additional Key Words and Phrases: Shared autonomy; Creativity; Storytelling; Child-Robot Interaction

1 INTRODUCTION

Creativity is an interaction that produces a novel and useful product, and it can be applied in many different contexts providing improvements in people's everyday lives [27]. Kovac [20] showed how creativity is highly correlated with prosocial behaviors, a sense of humor, and with less aggressive behaviors. Moreover, Plucker [26] presented how communication and creativity can positively impact conflict resolution and mediation.

Since it is not static and can be improved [27], creativity has been shown to increase with tools that help enhance it. These tools have been developed by psychologists and engineers, to practice divergent and convergent thinking and have been applied with varying degrees of success [7; 12; 15; 18; 24; 31]. To help children engage more when using these tools, robotic toys have been used and have been shown in past experiments to have a positive impact on children's creative processes, when programmed to stimulate that trait [1–3; 11].

These robotic toys were fully autonomous but shared autonomy allows a user to control the robot while the robot aids the user with autonomous behaviors. These behaviors have been shown to help the user increase efficiency in tasks, while still maintaining the user's trust in the system [16; 23]. These benefits from shared autonomy provide opportunities to explore them in creativity training tools.

Since creativity is not a static trait [27], it can also decrease, and this event is very noticeable in children at the age of 7 years old [27]. At this age, children have just joined the education system, and this pattern of decreased creativity can be related to the need of fitting in with their peers, by conforming to the group behaviors [27]. It can also be related to creativity being seen often as disruptive behaviors by school teachers [8]. This is a real problem and it is often referred to by the "Creativity Crises" [27] or the "Creativity Slump" [32].

Adding to that, with age increasing we get more cautious taking fewer risks which leads to less exploratory behaviors and less original ideas [27]. This pattern combined with the lack of tools to enhance creativity, and the ones that exist are very similar to tests [4], makes training and enhancing creativity a very hard and not-so-pleasing task.

In the area of children's creativity, storytelling has proven to be a valuable tool [4]. Storytelling provides the ability to create stories without imposed boundaries and guidelines that are present in other forms of story creation like writing. When robots are programmed to stimulate creativity, they can provide children with a playful and stimulating activity that helps enhance their creative processes [1–3; 11]. Since children suffer a decline in creativity at the age of 7 [27] robots can take a major part in helping increase children's creativity. These robots have features that regular toys do not usually have like programmable movements and rotations, programmable sounds, and different sorts of inputs (buttons and proximity sensors) that can also be programmed.

These features allow robotic toys to be used with shared autonomy, where a child controls the robot while telling a story, and the robot performs autonomous behaviors to trigger convergent or divergent thinking in the child. We study how shared autonomy helps children create stories, by combining the teleoperation of the robot with different emotion-related sounds, and emotion-related movements displayed by a robot during storytelling. The child using a controller, moves the robot constructing and telling the story at the same time, and the robot displays autonomous behaviors based on the content of the story.

We looked at the benefits of shared autonomy and storytelling and how we could combine them with the benefits of robots in creativity. This resulted in the first design of our system which was composed of a robot and a controller that could move the robot in discrete movements in four directions. This design was validated and suffered minor improvements in sounds and movements to allow better recognition of which emotions they were relating to.

These emotion-related sounds and movements were triggered by a wizard, that had an application where it was able to also track the total time of the experiment, pause the experiment when needed, and stop the experiment. The total amount of emotions available to trigger in the application was six: happy, sad, scared, disgust, surprise, and anger. All of these included movements and sounds.

We conducted a between-subjects experiment where a child was asked to create a story using the robot as the main character. To evaluate the impact of emotions displayed by the robot in the overall storytelling and in the child's creative process, we created three conditions: teleoperated where there were no emotions displayed during the storytelling, random where emotions were triggered in key moments of the story, but these emotions were picked randomly, and story where the emotions were triggered also in key moments of

Author's address: Miguel Azinheira, miguel.azinheira@tecnico.ulisboa.pt, Institute Superior Técnico, Lisboa, Portugal.

the story but were also related to the story itself. We then gathered the data from the creation of these stories and analyzed them.

2 RELATED WORK

In this section, we will cover the related work done in the area of storytelling with robots, robots, and creativity, and the related work done in shared autonomy.

2.1 Robots for Creativity

Creativity has been studied in social sciences for many decades on its benefits when trained and its drawbacks when neglected. Due to the many benefits of training creativity, it started to become a popular topic in the human-robot interactions field, with various papers covering the different aspects of the creative process and how robots can impact it positively.

Social robots have been shown to have a positive impact on creativity when compared to a PowerPoint presentation on a rock garden activity [17]. They have also been shown to increase children's creativity in a storytelling activity by stimulating convergent and divergent thinking, using the robot's movements [5]. When paired with emotional regulation techniques like promotion and prevention, social robots have also shown the ability to stimulate creativity using speech [11].

In a Droodle Creativity Game where participants have to generate creative titles for ambiguous images, robots when paired with curious expressions in body posture and eyes, have been shown to impact positively creativity [2]. When it comes to figural creativity, social robots have also been able to enhance it when programmed to stimulate that trait [3].

Social robots have also helped children solve a digital creative problem-solving game called Escape!Bot [1]. This game consisted of a sandbox-like game where players assembled contraptions to move the game's character to the goal. Children that interacted with the robot with a creative type of dialogue took significantly less time to win the game and did more spatial planning and used various unique objects as the winning strategy.

2.2 Storytelling with Robots

Chang et al. [9] studied how a robot could help children learn a second language. To accomplish this, they designed five modes for a robot to serve as a partner for teachers in the classroom. One of the modes was a storytelling mode that the robot would tell the stories using a male or female voice, perform comic actions and play sounds related to the story to increase engagement. The ability to make sounds and role-playing with different voices engaged more the children's attention.

Storytelling can also have benefits in pediatric rehabilitation, Plaisant et al. [25] conceived a storytelling robot for child patients that needed pediatric rehabilitation. The goal of the study was to create a robot to help children that needed to do rehabilitation exercises with lots of repetitions that children usually do not finish. By using the robot to create a story, children had to mimic the movements that the robot would do during the story. These movements were recorded as the child was doing them, and were used later in the story's construction. The authors found that children were very excited to use the prototype.

Munekata et al. [22] studied how robots could help autistic children or children with Down's syndrome in storytelling activities. The robot would stimulate children while they were writing an email by saying encouraging messages. The authors conducted the study for two years and collected data from the stories in the emails written by each child. The results show that the length of the stories increased with each session and that children started to produce more grammatically complex stories.

Sugimoto [30] created a system to enable the manipulation of a robot using a handheld projector in a storytelling activity. The robot supported successfully the children's storytelling activities and when interacting with the robot children engaged more in their tasks.

The novelty of robots can help capture children's attention and make them engage more in activities. Ribeiro et al. [28] used a Lego robotics kit to try to get more engagement from children in a storytelling activity and showed that students were able to construct and program the robot with high levels of motivation and enthusiasm.

2.3 Shared Autonomy in Human-Robot Interactions

Shared autonomy is the coordination between multiple agents on a task or system, where an agent remains autonomous, but respects the autonomy of other agents by adjusting its degree of autonomy [29]. To achieve shared autonomy, the agents should agree on common ground, be transparent about their goals, and communicate between them.

In assistive robots where a user teleoperates the robot, the input from the human and the autonomous assistance from the robot have to be combined to successfully reach a goal. Javdani et al. [16] proposed a solution to improve the shared autonomy between a robotic arm and the user input, by analyzing the history of user inputs and creating a distribution of user goals. This distribution is then used to predict the goal of the human based on the inputs, and provide a faster and more cost-effective solution to reach that goal. The authors conducted a study where participants were tasked with picking up an object using a teleoperated robotic arm using two different teleoperation systems. Each of the objects had to be picked up at least one time in any random order. To evaluate the efficacy of their system, the authors compared their method with a single-goal prediction method and found that their method helped users to finish their tasks faster and with less input provided.

User-driven customization can also be used to increase the efficiency and performance of assistive robots. Gopinath et al. [13] analyzed how user-driven customization performed against predefined optimization techniques, by conducting a study where the participant had to use a robotic arm in three different tasks. In the first task, the user was asked to use the robotic arm to reach a single object, using the system in full teleoperation mode and under three different predefined assistance levels. This task was designed to enable the user to get familiar with the control interface and the assistance system. The user would provide verbal feedback on the assistance level to a system operator who would then create a customized version and let the user test the system again. After the customization, the user was asked to perform two more tasks using two pre-defined assistance levels and the customized assistance level, and data was collected from these interactions. The results showed that when using the customized assistance level the users improved their performance on the tasks.

Nikolaidis et al. [23] tried to improve the performance of a human controlling a robotic arm by guiding the human in order to achieve an effective strategy, while still retaining his trust. To do this, the authors created an adaptation model where the system identified if the human controlling the robot was adaptable or stubborn. If the human was adaptable the robot would guide the human to achieve optimal performance in the task, but if the human was stubborn, the robot would comply with the human's orders to retain their trust. To test the system the authors conducted a study with three conditions: the robot using the adaptation model, the robot using a one-way adaptation model by always complying with the human, and the robot using a fixed policy with the objective of always reaching an optimal goal. After the activity, the authors conducted a questionnaire where the participants would rate their trust and perceived collaboration of the robot. The results showed that when using the one-way adaptation model the participants had the worst performance but would rate their trust in the robot higher. In the fixed policy model the participants would have the best performance of the three conditions but the worst levels of trust in the robot. In the adaptation model, the results were balanced between the one-way and the fixed policy, the participants had better performance than the participants in the one-way model and trusted the robot more than the participants in the fixed policy.

The teleoperation of the robot does not require the human to be at the same site as the robot, but using a remote teleoperation system presents vision challenges that impact the precision of the movements. Kofman et al. [19] created a non-contacting vision-based human-robot interface that allows the human to control a robot remotely while getting feedback on the robot's motion and environment. This interface was made using two cameras that tracked the hand of the human teleoperator and transmitted the movements parsed as commands to the robot on the remote site. After receiving the commands the robot applies them while also applying autonomous movements that help the operator. The authors conducted tests on the interface and found that the operator when using the interface was able to place an object on a target with high accuracy.

Shared autonomy has also been used to add some degree of autonomy to wheelchairs, allowing the user to reduce interactions during autonomous navigation. Chang et al. [10] created a wheelchair system that was able to navigate autonomously and choose the navigation path based on user preferences. The user would select a location and the system would try to calculate which path the user was more likely to choose. After this calculation, if the system was not able to predict the path the user had the option to choose one of the paths using a touch interface. This system only collected the user's preference from the path chosen in case of a tie in the calculation process, but there is valuable data that could be collected on user preferences based on previous manual navigations using the joystick or a simulated environment [21].

In shared autonomy systems, the interaction between the user and the robot has to be done using a controller that usually requires physical interaction from the user. Beraldo et al. [6] studied how a Brain-Computer Interface could be used by children to control a robot. This BCI allowed children to operate a robot by looking at images that would trigger brain signals, which were then interpreted and parsed as robot commands.

3 THE PROTOTYPE

We created a prototype to help children create stories, by stimulating their creative process, using a semi-autonomous robot that was teleoperated by a child and had autonomous behaviors triggered by a wizard. This prototype is composed of a mat with visual cues to help create the story, a robot that is controlled by a remote controller, and an app that allows the wizard to trigger the robot's actions. We tried to always keep the prototype simple so that it has fewer points of possible failures and that is easier to assemble.

3.1 Workspace

We chose a mat[1] that looked like a city map, that the robot could navigate through, with visual cues that would help to create the story. This mat had a length of 130 cm and a width of 133 cm, which gives it enough space for the robot to move around. It was also a thin mat with a thickness of 2 cm allowing it to be easily rolled and folded. These visual cues included a lighthouse surrounded by the sea with a boat and some sharks, a circus, a football field, a train track, a farm, a cave, and some other buildings like houses, schools, a gas station, and a train stop.

3.2 Robot

The Dash[2] robot was chosen to be used on this prototype since it provided out-of-the-box many features like real-time Bluetooth, programmable LEDs, potentiometers, and dual motors that allowed position precisely the robot's head, two powered wheels, and a set of speakers. It was also capable of storing sounds, which were recorded and uploaded on the manufacturer's provided app.

3.3 Controller

The remote chosen for the final prototype was a PlayStation 4[3] remote and was chosen because of the long-lasting battery and its proven durability. Since we wanted the child to be focused on the story, we chose the movements of the robot to be discrete. This would also allow us to save to a CSV file, each discrete movement the child did throughout the story as well as the direction of each movement (forward, backward, left, and right).

3.4 Behaviors

We developed multiple functions with a set of custom movements and sounds for each behavior of the robot. We created a list below to show the different sounds and movements that were made on each of the functions:

[1]https://www.ikea.com/pt/pt/p/stadsdel-tapete-30361910/
[2]https://www.makewonder.com/robots/dash/
[3]https://www.playstation.com/pt-pt/accessories/dualshock-4-wireless-controller/

- **Happy behavior:** The robot would play a happy sound twice followed by a 360 degrees rotation.
- **Sad behavior:** The robot would pitch its head down, play a sad sound twice, and pitch the head back to its starting position.
- **Scared behavior:** The robot would play a scared sound, rotate its head twice from right to left, play a scared sound again, and move back.
- **Disgusted behavior:** The robot would move back and play a disgusted sound.
- **Surprised behavior:** The robot would play a surprised sound, turn its entire body from right to left twice, with an interval of one second, and play a surprised sound again.
- **Angry behavior:** The robot would move forward, play an angry sound, turn its body from right to left, and play an angry sound again.

3.5 Shared Autonomy

The commands sent to the robot were made using an open-source library[4], that provided an abstraction layer for the hexadecimal values, which represented each possible action of the robot. We created a function for each of the behaviors that would follow a pattern between them. When a behavior was triggered, a command to change the robot's LEDs color to purple was sent to the robot, and when all the movements and sounds from the behavior were finished, a command to change the color of the robot's LEDs to blue was sent to the robot. This provided for the child, a visual cue signalizing that the robot was starting and ending an autonomous action.

We developed an app with a graphical user interface, that allowed the wizard to connect to the robot, and trigger the behaviors smoothly as the storytelling activity was happening. When the wizard initialized the app, a screen with three buttons regarding the different reactive strategies was presented. Choosing the Control option would make a timer appear and a button to start the activity. After pressing start, a CSV file was generated and opened, and a button with pause and stop would appear on the screen. The pause button was intended to pause the activity by disabling the remote, but keeping the current CSV file open, the stop button, on the other hand, was intended to stop the activity, by disabling the remote and also closing the file. The wizard also had a back button to return to the initial menu and these buttons were present on all three screens.

Choosing the Shared Autonomy Random option the same timer, timer controls, and back button as presented above, would appear, but it would also show a button that was able to trigger the different robot behaviors randomly. Choosing the Shared Autonomy Story option would show the timer and the other buttons present in the previous conditions, except the Random Behavior button, and it would also present six other buttons, that would allow the wizard to trigger the different behaviors.

The CSV generated after clicking start on all three screens, would have the same three columns: author, movement, and time. The author column had two options, child and wizard, which the first appeared when a child sent a command using the remote, and the second appeared when the wizard triggered a behavior. In the second column, movement, the options would vary depending on the first column. If in the first column the author was the child, then in the second column the only options available would be up, down, left, and right. If in the first column the option was wizard the only options available would be happy, sad, scared, disgusted, surprised, and angry.

[4]https://github.com/IsabelCanicoNeto/DashRobot

3.6 Architecture

The architecture behind our prototype was designed to keep the system simple and less prone to failure, and because of that, we opted for a robot that had all the desired features described above and it also had a couple of different libraries available to program it. The overall system was composed of a PlayStation 4 remote, a computer running a python script with a user interface, and the robot.

4 USER STUDY

The goal of our work was to investigate the impact of shared autonomy on children's creativity in a storytelling activity. We designed two types of reactive strategies: random and story. These reactive strategies were the type of autonomous behavior that the robot would display depending on the story contents. We also wanted to examine whether these different types of reactive strategies can impact the creative process differently. To reach that goal we conducted a between-subjects user study with three conditions: a random condition, where the robot would use the random reactive strategy and display random behaviors in key moments of the story; a story condition, where the robot would use the story reactive strategy and display behaviors in key moments of the story and based on the content of the story; and a control condition, where the robot would not display any emotions. In this chapter we will present our user study and its findings, ending with a discussion of the results.

4.1 Research Questions

The goal of our study was to answer these questions:

- **RQ1:** Does shared autonomy affect creativity in children during a storytelling activity?
- **RQ2:** Can different types of reaction strategies have a different impact on the creative process during storytelling?

4.2 Participants

The user study was conducted at two different places, an out-of-school study club - "Sala do Futuro", and a primary school - "Escola Básica Bolembre", both in Lisbon. Before starting, a consent form was handed to the parents of the children who wanted to participate in the study. After signing and giving consent, this form also allowed parents to choose if they want the face of their child hidden during the video recording. The user study was approved by the Ethics Committee of Instituto Superior Técnico.

In total, we had 54 children participating in the activity ranging from 5 to 13 years old, 23 females and 31 males. The average age of the participants was 9.130 years old (SD = 0.272).

4.3 Procedure

We chose to conduct a between-subjects study in order to compare the differences between the different types of reactive strategies of the robot. Each child participated in only one session, and the goal throughout the study was to have an equal amount of subjects per condition.

The sessions were conducted in a quiet room provided by the employees of the club and the school. Two researchers were present in all the sessions and were responsible for preparing the setup before the child arrived and guiding the child throughout the study. The session was divided into three main parts: the TCT-DP part, where the child was asked to complete a drawing, the activity part where the child was introduced to the activity and was asked to create a story using the robot, and the questionnaire part, where the child would finish the activity by responding to questions regarding the robot and the researchers.

Fig. 1. Study layout

In the first part of the activity, the child was introduced to both of the researchers, one of whom would guide the child to seat next to the mat and the robot (see Figure 1). The other researcher would hand the TCT-DP form A and a pen to the child, explaining how an artist started the drawing but was unable to finish it, and that we needed a complete drawing. At this time the robot was turned off to avoid distractions and the controller was hidden from view.

When the drawing was finished the researcher would explain, that the goal of the activity was to create a story, where the robot was the main character. Throughout the story, the child would be required to try to visit the different places on the mat, with the robot. Then the researcher would explain how the controller works by demonstrating how to move the robot up, down, left, and right and would hand the remote to the child. During this demonstration, the second researcher independently of the condition was sited on a chair with a laptop with the wizard software.

In the control condition, the robot was controlled by the child and would not display any behaviors. One of the researchers would guide the child to help avoid early withdrawals from the activity and would ask the child to give a name to the robot. The purpose of the researcher interacting with the child was also to encourage the child to tell the story out loud since the activity was being recorded for later coding. In the random and story conditions, the researcher would also introduce the activity the same way, and would not give any hints that the robot would do any sort of unexpected behavior. The difference between conditions was that in the random condition, the wizard, which had an app where it could trigger the actions of the robot, would click on the button random behavior when a key moment of the story was happening. These key moments were any emotions or actions that were relevant in the context of the story.

For the story condition, the key moments would be the same as for the random condition. The wizard would be sitting in the exact same spot as the other two conditions, with the laptop facing away from the child's face. The behaviors of the robot would be triggered by a set of buttons, that allow the robot to express 6 kinds of behaviors: happy, sad, scared, disgusted, surprised, and angry. The wizard was able to trigger these behaviors without being spotted because the laptop and the position of the researcher to the child would hide the hands of the researcher. Throughout the 54 subjects the wizard was always controlled by the same researcher, since key moments of the story can always be interpreted differently by listeners. Also, the researcher that was issuing the behaviors would avoid participating in guiding the child, to avoid bringing attention to the laptop.

When the child finished the story the researcher guiding the child would complement the story and retrieve the controller. The researcher controlling the wizard would hit stop to finish the data collection from the controller and the robot. At this point, the child was introduced to the last part of the activity where they had to respond to a questionnaire regarding the robot and the researchers. The researcher would hand a graph with 5 bars starting from completely disagree all the way to completely agree. This graph served to help children respond to verbal questions. For every 5 verbal questions, the child was asked to respond to a secret question by drawing on a piece of paper and inserting it into a box. When the questionnaire was finished the researchers would thank the child for their participation and guide the child back to the playground.

4.4 Measures and Data Analysis

The data collected from the activity was quantitative, being extracted from the CSV file export of the wizard app, combined with the data extracted from the video recordings of the interaction, from the TCT-DP test, and from the questionnaires at the end of the activity. To analyze all this data one of the researchers coded the results from the TCT-DP test, leading to a final result on the baseline creativity, and the second researcher coded the videos and audio from the sessions.

All the sessions were transcribed and coded in 4 metrics inspired by the work of Elgarf et al. [11]: fluency, flexibility, elaboration, and originality. Fluency was coded using a 0 to 5 score composed of the sum of the different story elements expressed throughout the story: character creation and description, describing the setting or location of the story, the plot, the conflict, and the resolution of the conflict. Fluency served to analyze the basic structure of the story since creativity should produce a perceptible product [27]. Flexibility was

Table 1. Levels of originality

Originality	Description
1	Very short and non-invested clauses about the locations present in the mat and repetitions of traditional story plots.
2	Creative ways of using the locations on the mat and surprising uses of characters and actions that can have a background of traditional story plots.
3	Creation of new locations not present on the mat, acting and first-person narrative, and rare story plots.

coded by counting the total different number of characters, actions, scenarios, and affective expressions present in the story. These characters, actions, and scenarios would only count the first time they appeared in the story and repetitions of these were ignored. Elaboration was coded using the videos and audio from the sessions and consisted of the number of seconds the child was speaking about the story. The last metric coded from the transcriptions was originality, which consisted of the overall originality of the story measured from 1 to 3, where 1 was low originality and 3 was high originality, using the coding scheme present in Table 1.

From the wizard app, we extracted from the CSV the total time interacting with the robot, the total number of movements issued with the remote by the child, and the number of interactions from the wizard. These interactions from the wizard consisted of the 6 behaviors displayed by the robot.

Regarding the questionnaire, we extracted data related to the researchers by asking questions to ensure the researchers were nice and that the activity was well explained. We also extracted metrics related to the perception of the robot by the child and the overall perception of the activity itself.

4.5 Findings

The data gathered from the activity was collected and we performed a statistical analysis to try and answer our research questions. We present the findings in this section.

4.5.1 Baseline Creativity. **Baseline creativity levels were not different between the conditions.** To analyze the baseline creativity levels among children we compared the different results of the TCT-DP tests. We expected the levels of creativity between the condition's groups, not to be statistically significant. From our analysis, we found no statistically significant difference between control ($M = 21.667, SD = 8.568$), random ($M = 22.833, SD = 9.811$), and story ($M = 19.944, SD = 7.780$), as determined by one-way ANOVA ($F(2, 51) = 0.495, p = 0.612$). This goes in line with what we expected, confirming that neither of the condition's groups started with more or less creative children than the others.

4.5.2 Story Creation. **The difference in time interacting with the robot was not statistically significant between conditions.** This metric was retrieved from the CSV file and since it wasn't normally distributed we conducted a Kruskal-Wallis Test. We were expecting that the total amount of time could be bigger in the random and story conditions but from our analysis, this did not happen. The test showed that there was no statistically significant difference in the time spent storytelling with the robot between control ($M = 303.444, SD = 270.667$), random ($M = 321.944, SD = 344.677$), and story ($M = 236.500, SD = 171.783$) groups $X^2(2) = 0.846, p = 0.655$ with a mean rank duration of 30.28 for control, 26.28 for random and 25.94 for story.

One of our concerns initially was that the researcher that was operating the wizard would perform a different amount of emotions between the random and story conditions. **The number of emotions triggered by the wizard was not different between the random and story conditions.** We conducted a Mann-Whitney U Test and found no significant differences between random ($M = 9.667, SD = 8.275$) and story ($M = 6.111, SD = 3.428$) conditions ($U = 129.5, p = 0.301$).

The number of instructions sent by each child from the controller during the activity was not different between conditions. The results from the Kruskal-Wallis Test show that the number of key presses on the remote was not statistically significantly different between control ($M = 206.611, SD = 135.061$), random ($M = 278.389, SD = 222.818$), and story ($M = 193.667, SD = 138.924$) conditions $X^2(2) = 0.680, p = 0.712$. We can also observe that the mean rank of the random condition was 29.92, the mean rank of the control condition was 26.83, and the mean rank of the story condition was 25.75. This shows how the interactions of the robot in the random and story condition were not affecting the total number of movements issued by the child. This is both positive and negative because we would hope that the agency from the robot would stimulate the child to make more movements but at the same time, it also presents that the robot was not affecting negatively the number of movements issued by the child.

The TCT-DP test was significantly and positively correlated with fluency, flexibility, and elaboration. We ran a Spearman's rank-order correlation and found a strong, positive correlation between the metrics of fluency ($M = 3.815, SD = 1.375$) ($r_s(54) = 0.383, p = 0.004$), flexibility ($M = 39.259, SD = 37.484$) ($r_s(54) = 0.397, p = 0.003$), and elaboration ($M = 136.390, SD = 138.768$) ($r_s(54) = 0.459, p = 0.001$) with the TCT-DP test results, but we did not found a correlation between originality ($M = 1.926, SD = 0.797$) and the TCT-DP test results ($M = 21.482, SD = 8.676$) ($r_s(54) = 0.262, p = 0.055$).

With the activity having no time limit some of the metrics could increase with time. Knowing this we normalized the story metrics, fluency ($M = 0.046, SD = 0.035$), flexibility ($M = 0.318, SD = 0.104$), and originality ($M = 0.022, SD = 0.016$), by the story elaboration. We then ran a Kruskal-Wallis Test and found no statistically significant differences between conditions with the metrics of fluency ($X^2(2) = 3.573, p = 0.168$), flexibility ($X^2(2) = 1.402, p = 0.496$), and originality normalized ($X^2(2) = 4.225, p = 0.121$).

4.5.3 Analysis Clustered by the Initial Creativity Levels. One of the goals of this study was to find out if the robot had any impact on children's creativity during storytelling, so we had to compare the baseline of the levels of creativity to the levels of creativity during storytelling. Since we used 3 levels of originality, one of the most important metrics of creativity [27], we decided to create 3 clusters from the baseline levels of creativity (TCT-DP test results): a cluster for the test results with low creativity ($N = 14$) with an average age of 8.786 years old ($SD = 1.762$), a cluster for the test results with average creativity ($N = 20$) with an average age of 8.450 years old

($SD = 1.932$), and a cluster for the test results with high creativity ($N = 20$) with an average age of 10.050 years old ($SD = 1.959$). The low creativity cluster had 4 participants in the control and random conditions and 6 participants in the story condition. The average creativity cluster had 8 participants in the control condition, 6 participants in the random condition, and 6 participants in the story condition. The high creativity cluster had 6 participants in the control condition, 8 in the random condition, and 6 in the story condition.

Using these clusters we ran a Kruskal-Wallis Test and found that in the low baseline creativity cluster, there was a statistically significant difference in fluency normalized by the elaboration ($M = 0.068, SD = 0.050$) (time spent telling the story), between the different conditions $X^2(2) = 6.157, p = 0.046$ and that there was also a statistically significant difference in originality normalized by the elaboration ($M = 0.032, SD = 0.024$), between the different conditions $X^2(2) = 6.019, p = 0.049$. However, the flexibility normalized by the elaboration ($M = 0.387, SD = 0.152$) was not statistically significantly different between conditions $X^2(2) = 0.424, p = 0.809$.

On the medium creativity cluster, we didn't find any significant differences between conditions for fluency ($M = 0.038, SD = 0.024$) $X^2(2) = 3.904, p = 0.142$, flexibility ($M = 0.292, SD = 0.077$) $X^2(2) = 1.205, p = 0.548$, and originality ($M = 0.019, SD = 0.011$) $X^2(2) = 2.921, p = 0.232$, when normalized by the elaboration. As for the children in the high creativity cluster we also didn't find any significant differences between conditions for fluency ($M = 0.039, SD = 0.024$) $X^2(2) = 0.262, p = 0.877$, flexibility ($M = 0.296, SD = 0.058$) $X^2(2) = 0.123, p = 0.941$, and originality ($M = 0.018, SD = 0.010$) $X^2(2) = 0.123, p = 0.941$, also when normalized by the elaboration. **With these results we can see that children with lower baseline creativity levels benefited from the behaviors of the robot.**

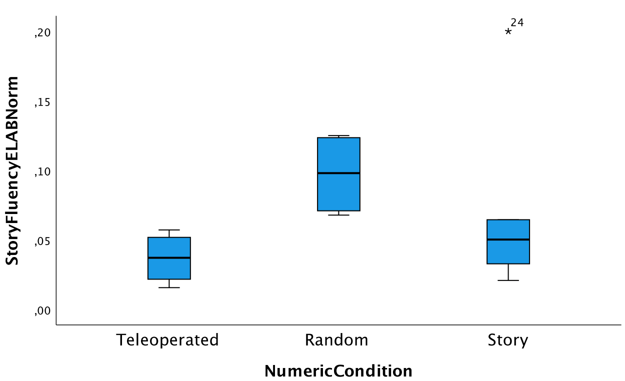

Fig. 2. Fluency when normalized by elaboration on low creativity baseline levels

To identify which of the conditions had a significant difference from the others on the low creativity cluster, on story fluency when normalized by the elaboration, we ran three sets of Mann-Whitney U tests to compare them. We found that the teleoperated condition ($M = 0.037, SD = 0.019$) was not statistically significantly different from the story condition ($M = 0.070, SD = 0.066$) ($U = 7, p = 0.286$) but was statistically significantly lower than the random condition ($M = 0.097, SD = 0.030$) ($U = 0, p = 0.021$). When comparing the story and random conditions on the story fluency, when normalized by the elaboration, we found no statistically significant difference between the two ($U = 4, p = 0.088$). **The results from these tests show us that children with lower baseline creativity levels, benefited more in the fluency of the stories, from the behaviors of the robot when using the random condition.**

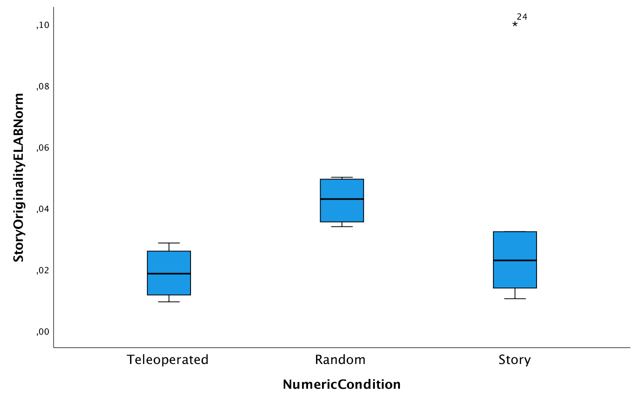

Fig. 3. Originality when normalized by elaboration on low creativity baseline levels

We also analyzed which of the conditions had a significant difference from the others, on story originality when normalized by the elaboration. We ran three sets of Mann-Whitney U tests to compare them on the cluster of low baseline creativity. From the results of the tests, we found that the teleoperated condition ($M = 0.019, SD = 0.009$) was not statistically significantly different from the story condition ($M = 0.034, SD = 0.034$) ($U = 7.5, p = 0.336$) but was statistically significantly lower than the random condition ($M = 0.042, SD = 0.008$) ($U = 0, p = 0.021$). When comparing the story and random conditions on the story originality, when normalized by the elaboration, we found no statistically significant difference between the two ($U = 4, p = 0.088$). **The results from these tests show us that children with lower baseline creativity levels, benefited more in the originality of the stories, from the behaviors of the robot when using the random condition.**

4.5.4 Questionnaire Findings. **Children perceived equally the investigators between conditions.** We ran multiple Kruskal-Wallis Tests that showed that there was no statistically significant difference in how nice the researchers were between control ($M = 4.882, SD = 0.485$), random ($M = 4.778, SD = 0.428$) and story ($M = 4.824, SD = 0.393$) conditions $X^2(2) = 1.572, p = 0.456$ and that there was no statistically significant difference in how the researcher explained the activity between conditions $X^2(2) = 0.206, p = 0.902$ with control with an average score of 4.824 ($SD = 0.529$), random with an average score of 4.778 ($SD = 0.548$), and story with an average score of 4.824 ($SD = 0.529$).

From the questionnaires, we also tried to perceive if there was any significant difference in children liking the activity between the conditions. **Children liked to do the story with the robot equally**

between conditions. We ran a Kruskal-Wallis Test that showed no difference in likability between control ($M = 4.706, SD = 0.470$), random ($M = 4.222, SD = 1.166$), and story ($M = 4.647, SD = 0.606$) conditions $X^2(2) = 1.417, p = 0.492$. This was common, children that weren't participating in the activity were excited to participate, and children that finished the activity would talk positively about the activity with their peers.

We expected that children would rate the robot more intelligently when the robot was in story condition. This was not the case, we ran a Kruskal-Wallis Test and found no statistically significant difference between conditions $X^2(2) = 1.139, p = 0.566$ with control with an average score of 4.324 ($SD = 0.983$), random with an average score of 3.778 ($SD = 1.478$), and story with an average score of 4.353 ($SD = 0.702$). **Children rated equally the robot's intelligence.**

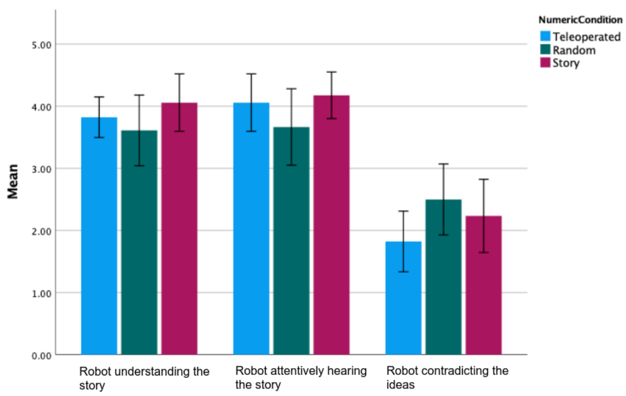

Fig. 4. Answers on the perception of the robot during the storytelling

Children's perspectives on the robot were also important to understand. We ran a set of Kruskal-Wallis Tests and found no statistically significant difference between conditions on the question about the robot understanding their story ($M = 3.827, SD = 0.923$) $X^2(2) = 1.755, p = 0.416$ with a mean rank of 25.320 for control (N=17), 24.140 for random (N=18), and 30.180 for story (N=17), no statistically significant difference between conditions on the question about the robot hearing attentively their story ($M = 3.962, SD = 0.989$) $X^2(2) = 1.353, p = 0.508$ with a mean rank of 27.500 for control (N=17), 23.440 for random (N=18), and 28.740 for story (N=17), and no statistically significant difference between conditions on the question about the robot contradicting their ideas ($M = 2.192, SD = 1.103$) $X^2(2) = 3.276, p = 0.194$ with a mean rank of 21.680 for control (N=17), 30.530 for random (N=18), and 27.060 for story (N=17). Even though this data is not conclusive, we can see a clear pattern in Figure 4. **Children rated equally the robot understanding their stories, hearing them attentively, and contradicting their ideas.**

We also analyzed the perception of children on the robot's autonomy, anthropomorphism, animacy, trust, closeness, similarity, and helpfulness on a scale of 1 to 5. The average results for autonomy in the control condition were 2.589 ($SD = 0.976$), in the random condition 2.870 ($SD = 1.011$), and in the story condition 3.275 ($SD = 0.775$). For anthropomorphism, the average results in the control condition were 3.397 ($SD = 0.923$), in the random condition 3.000 ($SD = 0.947$), and in the story condition 3.603 ($SD = 0.673$). As for animacy, the average results in the control condition were 2.784 ($SD = 1.040$), in the random condition 2.778 ($SD = 1.120$), and in the story condition 3.098 ($SD = 0.941$).

The average results for trust in the control condition were 4.216 ($SD = 0.716$), in the random condition 3.833 ($SD = 0.752$), and in the story condition 4.137 ($SD = 0.472$). For closeness, the average results in the control condition were 4.221 ($SD = 0.637$), in the random condition 4.097 ($SD = 0.713$), and in the story condition 4.118 ($SD = 0.662$). As for similarity, the average results in the control condition were 2.961 ($SD = 1.053$), in the random condition 2.389 ($SD = 0.842$), and in the story condition 2.843 ($SD = 0.657$). The last metric we analyzed was helpfulness with an average of 4.589 ($SD = 0.507$) for the control condition, 4.167 ($SD = 1.150$) for the random condition, and 4.412 ($SD = 0.618$) for the story condition.

We ran a set of Kruskal-Wallis Tests and didn't find any statistically significant difference between conditions on autonomy $X^2(2) = 4.145, p = 0.126$, trust $X^2(2) = 3.221, p = 0.200$ or helpfulness $X^2(2) = 0.976, p = 0.614$. We also didn't find any statistically significant difference between conditions on anthropomorphism ($F(2, 49) = 2.239, p = 0.117$), animacy ($F(2, 49) = 0.534, p = 0.590$), closeness ($F(2, 49) = 0.167, p = 0.847$), or similarity ($F(2, 49) = 2.145, p = 0.128$) using the one-way ANOVA.

4.6 Discussion

At the beginning of this chapter, we defined two research questions that we ought to respond to with our collected data and our analysis of that data. We present in this section the responses to our initial research questions.

RQ1: Does shared autonomy affect creativity in children during a storytelling activity?

During the storytelling activity, we collected different types of data regarding the robot's movements and the story itself. To compare the differences in creativity levels, we had to compare to a baseline level of creativity which was provided by the pre-test. We found out that the baseline creativity test was highly correlated with a couple of metrics extracted from the story (fluency, flexibility, and elaboration), which gave us the confidence to compare the results from the pre-test to the metrics collect during the storytelling.

We created three clusters containing children with low, average, and high levels of creativity and compared the different metrics. For the average and high levels of creativity clusters, we did not find any statistically significant differences between conditions in the results. However, for the low level of creativity cluster, we found statistically significant differences between conditions for fluency and originality.

These results could be explained by looking at the different factors that affected the outcome of the activity. For instance, we can see that each condition on each cluster only had a maximum of 8 participants, and this is a relatively small sample since the differences between groups can sometimes be very small. We didn't also enforce a minimum time length for a story which meant that children with shorter stories from the shared-autonomy conditions had almost no actions triggered by the wizard.

To give a final answer to our research question, we can say that shared autonomy has a direct impact on the creative process of a storytelling activity in children with low creativity levels when using the random reactive strategy.

RQ2: Can different types of reactive strategies have a different impact on the creative process during storytelling?

We found out that when using the random behavior, the story fluency and originality suffered a significant positive impact on children from the low creativity cluster. This could be explained by the fact that when using the random reactive strategy, the robot would display emotions not related to the story, sparking more divergent thinking in children that were less creative from the start.

The data collected from the questionnaires didn't provide any conclusive data, on the impact of the different types of reactive strategies. Even though no statistically significant data was found, we can see a pattern between the two. Children seem to think that the robot understands better their story when using the story condition and that the robot contradicts more their story on the random condition. This pattern that we can see in Figure 4, gives us a hint that the questionnaires used were possibly not the best suited for children since children often try to not give low grades on answers [14].

To give a final answer to this research question, we found that in children with lower baseline creativity levels, the random reactive strategy had a significant positive impact on the creative process in a storytelling activity (fluency and originality) when compared to the story reactive strategy. We also found that for children with average, and high creativity levels the reactive strategies despite not having a positive impact also did not have a negative impact.

4.7 Limitations and Future Work

Our study had some limitations and because of that, it might have impacted negatively our results. We conducted a between-subjects study with three groups with 18 children per condition. When we created the three clusters, each condition was left with a small number of children. To collect more valuable data we would have to either increase the number of participants or decrease the number of conditions. Time was also a big factor that we did not take into account when we designed the study. We were not expecting the disparities in story lengths, with some children making very short stories and others very long stories. In the shorter stories, the wizard was not able to trigger the necessary number of behaviors to have a clear impact on the story outcome, which had a negative impact on our collected data. One improvement to take into account in future studies would be to create a maximum and minimum story size. This could be enforced by creating a set of tasks parallel to the story, that requires the robot to navigate through the different places on the mat.

In our study, we used the key moments of the story to trigger the robot's behaviors but these were sometimes very rare and even on the story condition, the behaviors of the robot were not suitable for certain key moments of the story. For example, a child communicating that the robot is playing a football game leaves the wizard to choose a behavior that would best fit the story, but the wizard's interpretation could differ from the child's interpretation and the emotion triggered can seem unrelated to the child. In future work with the suggestion of a fixed time span for the activity, the behaviors of the robot could be triggered based on time instead of the key moments from the story. That would allow children with lower descriptive stories to be exposed to the same amount of behaviors as children with more descriptive stories. In a future study, it could also be interesting to try and increase the time of each behavior to try and catch better the attention of the child.

The behaviors were also too simple, more complex behaviors could have a different impact on creativity. In future studies, speech could be added to better express the behavior of the robot, and more expressive movements could also be explored.

5 CONCLUSION

In this paper, we presented a complete system to measure the impact of simulated shared autonomy by a wizard, on the creative process during a storytelling activity. The system was robust and easy to use and made the activity easy and pleasing to children. It also provided solid data on the interactions of the child and the wizard with the robot.

We conducted a between-subjects study on 54 participants that aimed to compare the impact of shared autonomy between 3 different conditions: teleoperated where the robot was solely controlled by the child, random where the robot was controlled by the child and had random behaviors triggered by the wizard at key moments of the story to simulate semi-autonomy, and story where the robot was also controlled by the child but displayed behaviors that were also triggered by the wizard, that would better fit in the story at its key moments. Results show that randomizing the emotions during the storytelling activity increased fluency and originality in children with lower baseline creativity scores. As the first work combining the topics of shared autonomy, storytelling, and creativity, it provided valuable insights into the area. It also provided concrete data on how the different types of reactive strategies in a simulated semi-autonomy environment, have impacted the different aspects of the storytelling activity. Despite these promising results further studies should be conducted to evaluate them in the longer term.

REFERENCES

[1] Ali, S., Devasia, N. E., and Breazeal, C. Escape! bot: Social robots as creative problem-solving partners. In *Creativity and Cognition* (2022), pp. 275–283.

[2] Ali, S., Moroso, T., and Breazeal, C. Can children learn creativity from a social robot? In *Proceedings of the 2019 on Creativity and Cognition.* 2019, pp. 359–368.

[3] Ali, S., Park, H. W., and Breazeal, C. A social robot's influence on children's figural creativity during gameplay. *International Journal of Child-Computer Interaction 28* (2021), 100234.

[4] Alves-Oliveira, P., Arriaga, P., Cronin, M. A., and Paiva, A. Creativity encounters between children and robots. In *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2020), IEEE, pp. 379–388.

[5] Alves-Oliveira, P., Arriaga, P., Paiva, A., and Hoffman, G. Yolo, a robot for creativity: A co-design study with children. In *Proceedings of the 2017 Conference on Interaction Design and Children* (2017), pp. 423–429.

[6] Beraldo, G., Tortora, S., and Menegatti, E. Towards a brain-robot interface for children. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (2019), IEEE, pp. 2799–2805.

[7] Birdi, K., Leach, D., and Magadley, W. Evaluating the impact of triz creativity training: an organizational field study. *R&D Management 42*, 4 (2012), 315–326.
[8] Chan, D. W., and Chan, L.-K. Implicit theories of creativity: Teachers' perception of student characteristics in hong kong. *Creativity Research Journal 12*, 3 (1999), 185–195.
[9] Chang, C.-W., Lee, J.-H., Chao, P.-Y., Wang, C.-Y., and Chen, G.-D. Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Journal of Educational Technology & Society 13*, 2 (2010), 13–24.
[10] Chang, Y., Kutbi, M., Agadakos, N., Sun, B., and Mordohai, P. A shared autonomy approach for wheelchair navigation based on learned user preferences. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* (2017), pp. 1490–1499.
[11] Elgarf, M., Calvo-Barajas, N., Alves-Oliveira, P., Perugia, G., Castellano, G., Peters, C., and Paiva, A. " and then what happens?" promoting children's verbal creativity using a robot. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction* (2022), pp. 71–79.
[12] Fink, A., Koschutnig, K., Benedek, M., Reishofer, G., Ischebeck, A., Weiss, E. M., and Ebner, F. Stimulating creativity via the exposure to other people's ideas. *Human brain mapping 33*, 11 (2012), 2603–2610.
[13] Gopinath, D., Jain, S., and Argall, B. D. Human-in-the-loop optimization of shared autonomy in assistive robotics. *IEEE Robotics and Automation Letters 2*, 1 (2016), 247–254.
[14] Hall, L., Hume, C., and Tazzyman, S. Five degrees of happiness: Effective smiley face likert scales for evaluating with children. In *Proceedings of the the 15th international conference on interaction design and children* (2016), pp. 311–321.
[15] Howard, T. J., Culley, S. J., and Dekoninck, E. A. Stimulating creativity: A more practical alternative to triz. In *DS 58-5: Proceedings of ICED 09, the 17th International Conference on Engineering Design, Vol. 5, Design Methods and Tools (pt. 1), Palo Alto, CA, USA, 24.-27.08. 2009* (2009).
[16] Javdani, S., Srinivasa, S. S., and Bagnell, J. A. Shared autonomy via hindsight optimization. *Robotics science and systems: online proceedings 2015* (2015).
[17] Kahn, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Shen, S., Ruckert, J. H., and Gary, H. E. Human creativity can be facilitated through interacting with a social robot. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2016), IEEE, pp. 173–180.
[18] Kim, Y., and Park, N. The effect of steam education on elementary school student's creativity improvement. In *Computer applications for security, control and system engineering*. Springer, 2012, pp. 115–121.
[19] Kofman, J., Wu, X., Luu, T. J., and Verma, S. Teleoperation of a robot manipulator using a vision-based human-robot interface. *IEEE transactions on industrial electronics 52*, 5 (2005), 1206–1219.
[20] Kováč, T. Creativity and prosocial behavior. *Studia Psychologica 40*, 4 (1998), 326–330.
[21] Kutbi, M., Chang, Y., Sun, B., and Mordohai, P. Learning to navigate robotic wheelchairs from demonstration: Is training in simulation viable? In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops* (2019), pp. 0–0.
[22] MUNEKATA, T., Fujita, Y., and Nishizawa, T. New learning environment for enhancing storytelling activities of children with intellectual disabilities/autism using a personal robot in the classroom. *Semantic Scholar* (2009).
[23] Nikolaidis, S., Zhu, Y. X., Hsu, D., and Srinivasa, S. Human-robot mutual adaptation in shared autonomy. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI* (2017), IEEE, pp. 294–302.
[24] O'Donoghue, D., Abgaz, Y., Hurley, D., and Ronzano, F. Stimulating and simulating creativity with dr inventor.
[25] Plaisant, C., Druin, A., Lathan, C., Dakhane, K., Edwards, K., Vice, J. M., and Montemayor, J. A storytelling robot for pediatric rehabilitation. In *Proceedings of the fourth international ACM conference on Assistive technologies* (2000), pp. 50–55.
[26] Plucker, J. A. Positive approaches to preventing school violence: Peace building in schools and communities, 2000.
[27] Plucker, J. A., Beghetto, R. A., and Dow, G. T. Why isn't creativity more important to educational psychologists? potentials, pitfalls, and future directions in creativity research. *Educational psychologist 39*, 2 (2004), 83–96.
[28] Ribeiro, C. R., Coutinho, C. P., and Costa, M. F. Robotics in child storytelling.
[29] Schilling, M., Burgard, W., Muelling, K., Wrede, B., and Ritter, H. Shared autonomy—learning of joint action and human-robot collaboration, 2019.
[30] Sugimoto, M. A mobile mixed-reality environment for children's storytelling using a handheld projector and a robot. *IEEE Transactions on Learning Technologies 4*, 3 (2011), 249–260.
[31] Tahirsylaj, A. S. Stimulating creativity and innovation through intelligent fast failure. *Thinking Skills and Creativity 7*, 3 (2012), 265–270.
[32] Torrance, E. P. A longitudinal examination of the fourth grade slump in creativity. *Gifted Child Quarterly 12*, 4 (1968), 195–199.